OM nucleic - nucleic search, using sw model

```
           GenCore version 5.1.6
           Copyright (c) 1993 - 2004   Compugen Ltd.

Run on:          May  5, 2004, 09:55:26 ; Search time 4891.91 Seconds
                                          (without alignments)
                                          11571.357 Million cell updates/sec

Title:           US-09-901-572A-1
Perfect score:   1306
Sequence:        1 aaaaacatcagattgttaat........taaaatcgtttatcaggtt 1306

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters:      6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  GenEmbl:*
    1:  gb_ba:*
    2:  gb_htg:*
    3:  gb_in:*
    4:  gb_om:*
    5:  gb_ov:*
    6:  gb_pat:*
    7:  gb_ph:*
    8:  gb_pl:*
    9:  gb_pr:*
   10:  gb_ro:*
   11:  gb_sts:*
   12:  gb_sy:*
   13:  gb_un:*
   14:  gb_vi:*
   15:  em_ba:*
   16:  em_fun:*
   17:  em_hum:*
   18:  em_in:*
   19:  em_mu:*
   20:  em_om:*
   21:  em_or:*
   22:  em_ov:*
   23:  em_pat:*
   24:  em_ph:*
   25:  em_pl:*
   26:  em_ro:*
   27:  em_sts:*
   28:  em_un:*
   29:  em_vi:*
   30:  em_htg_hum:*
   31:  em_htg_inv:*
   32:  em_htg_other:*
   33:  em_htg_mus:*
   34:  em_htg_pln:*
   35:  em_htg_rod:*
   36:  em_htg_mam:*
   37:  em_htg_vrt:*
   38:  em_sy:*
   39:  em_htgo_hum:*
   40:  em_htgo_mus:*
   41:  em_htgo_other:*
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1280.4	98.0	2144	6	AR035279	AR035279 Sequence
	2	1270	97.2	1387	6	AR035276	AR035276 Sequence
	3	1264	96.8	1305	6	I17387	I17387 Sequence 1
	4	1262.4	96.7	1387	6	E09301	E09301 DNA encodin
	5	1179.4	90.3	302070	1	AB016968	AB016968 Mycoplasm
	6	1086.8	83.2	2014	6	AR035278	AR035278 Sequence
	7	1028	78.7	1152	6	AX665164	AX665164 Sequence
	8	989.4	75.8	1015	6	E09896	E09896 Mycoplasma
	9	986.2	75.5	1015	6	S65869	S65869 TM-1=29 kda
	10	921.6	70.6	1082	6	AX665187	AX665187 Sequence
	11	824.2	63.1	853	6	E02348	E02348 DNA sequenc
	12	679.8	52.1	708	6	E02342	E02342 DNA sequenc
	13	250.6	19.2	301042	1	AB016967	AB016967 Mycoplasm
	14	249	19.1	4568	1	AF210770	AF210770 Mycoplasm
c	15	222.8	17.1	301901	1	AB016969	AB016969 Mycoplasm
	16	217	16.6	7434	1	AF275312	AF275312 Mycoplasm
	17	216.6	16.6	2369	6	AR035275	AR035275 Sequence
	18	216.6	16.6	2369	6	E09895	E09895 Mycoplasma
	19	215	16.5	3331	1	AF141940	AF141940 Mycoplasma
	20	206.2	15.8	9720	1	MYCPMGAB	L28424 Mycoplasma
c	21	203.8	15.6	302070	1	AB016968	AB016968 Mycoplasm
	22	199.8	15.3	3894	1	AY065985	AY065985 Mycoplasm
	23	193.4	14.8	4715	1	AF032890	AF032890 Mycoplasm
	24	191.8	14.7	8904	1	MGU90714	U90714 Mycoplasma
	25	190.2	14.6	3897	1	AF053978	AF053978 Mycoplasm
	26	173.2	13.3	2803	1	MYCHAHOM	M83178 Mycoplasma
	27	173.2	13.3	2809	1	S55216	S55216 pMGA1.2=hem
	28	136.4	10.4	702	6	E02344	E02344 DNA sequenc
c	29	92	7.0	170627	2	AC125567	AC125567 Rattus no
c	30	89	6.8	146275	2	AL935272	AL935272 Danio rer
c	31	88.2	6.8	116807	3	BX890614	BX890614 Danio rer
c	32	85	6.5	110000	3	AC116305_1	Continuation (2 of
c	33	84.6	6.5	198743	2	BX530070	BX530070 Danio rer
	34	82.4	6.3	94534	5	AL929250	AL929250 Zebrafish
c	35	82.2	6.3	155204	2	AC007926	AC007926 Trypanoso
	36	80	6.1	154071	3	AC115598	AC115598 Dictyoste
	37	79.2	6.1	153751	2	AC116551	AC116551 Dictyoste
	38	79.2	6.1	225581	2	BX537105	BX537105 Danio rer
	39	78.8	6.0	199347	2	BX321884	BX321884 Danio rer
	40	78.6	6.0	213544	2	BX510939	BX510939 Danio rer
	41	78.2	6.0	594	6	E02345	E02345 DNA sequenc
	42	78.2	6.0	183357	2	BX569779	BX569779 Danio rer
	43	77.6	5.9	10593	3	AY232271	AY232271 Dictyoste
	44	77.4	5.9	335050	3	PFA929356	PFA929356 Plasmodiu
c	45	77.2	5.9	67970	3	PFMAL1P3	AL031746 Plasmodiu

ALIGNMENTS

```
RESULT 1
AR035279
LOCUS       AR035279       2144 bp    DNA    linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5871742.
ACCESSION   AR035279
VERSION     AR035279.1  GI:5951947
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2144)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
```

Pred. No. is the number of results predicted by chance to have a

JOURNAL   Patent: US 5871742-A 9 16-FEB-1999;
FEATURES
     source   1..2144
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   98.0%; Score 1280.4; DB 6; Length 2144;
Best Local Similarity   98.8%; Pred. No. 9.2e-177;
Matches 1290; Conservative   0; Mismatches 16; Indels   0; Gaps   0;

Qy   1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
            ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy   61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAAATGGTTTTCTTATC 120
            ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAAATGGTTTTCTTATC 120

Qy   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGAT 180
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGAT 180

Qy   181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy   241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300

Qy   301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAGAGCAGCGGCGAATGGAGTTAACAGAT 360
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCAGCGGCGAATGGAGTTAACTGAT 360

Qy   361 CTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACTACAAGAGAATGCCAAGATTGAAGCT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACTACAAGAGAATGCTAAGATTGAAGCT 420

Qy   421 AGTTTATCATCTGCTTATAGTGAAGTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGTTTATCATCTGCTTATAGTGAAGTGAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy   481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCTAATCAACCAAGCTAATACGGAT 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCTAATCAACCAAGCTAATACGGAT 540

Qy   541 AAAACGACTTTTGATAATGACACCCAAATTTAGTTGAAGCATACAAAGACTAAAAACC 600
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGACTAAAAACC 600

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCACAACTGCTTATAATCAAATT 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCAACTGCTTATAATCAGATT 660

Qy   661 CGCAATAATTAGTGGATCTATACAAATAAGTTAGTTGTTAATACTAAAACACTAGAT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGTAATAATTAGTGGATCTATACAAATAAGTTAGTTGTTTAATACTAAAACACTAGAT 720

Qy   721 CCACTAAATGGGGGAACGCTTTTGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCACTAAATGGGGGAATGCTTTTGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy   781 AATAATACGTTATCAACTATTAATTAATGAACAACAAAGACTAATGCATTATCTAATAGT 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AATAATACGTTATCAACTATTAATTAATGAACAACAAAGACTAATGCATTATCTAATAGT 840

Qy   841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAAT 960
            ||||||||||||||||||||||||
Db   901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAAT 960

Qy   961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAGAACCTTCAAGTAGAATTCTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TATAAATATCCAAGAAGGACCGTTTGGAATGGTGATGAGAACCTTCAAGTAGAATTCTGCA 1020

Qy   1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTTATAGTTAGTGGAACAAACACGAAG 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTTATAGTTAGTGGAACGAACGAAG 1080

Qy   1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
            ||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db   1081 TACCAATTTAGTTTTTAGCAACTACTGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140

Qy   1141 TTGGTTAAAGCAGTGATGCTAATAACGTTGGATTACAATACAATAATAATGGAAAT 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGTTAAAGCAGTGATGCTAATAACGTTGGATTACAATACAATACAATAATAATGGAAAT 1200

Qy   1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306

RESULT 2
AR035276
LOCUS   AR035276   1387 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5871742.
ACCESSION   AR035276
VERSION   AR035276.1 GI:5951944
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1387)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE   Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 3 16-FEB-1999;
FEATURES   Location/Qualifiers
     source   1..1387
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   97.2%; Score 1270; DB 6; Length 1387;
Best Local Similarity   98.8%; Pred. No. 3.4e-175;
Matches 1290; Conservative   0; Mismatches 15; Indels   1; Gaps   1;

Qy   1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy   61 AATCCTAAATAAATAAGCCGTTAATCGTAAAATCGATATCTTAACTAACTAAAAACACTAGAT 120
            ||||||||||||||||||||||| ||||||||||||||| |||||||||||||||||||
Db   61 AATCCTAAATAAATAAGCCGTTAATCGTAAATCGATATCTTAACTAACTAAAAACACTAGAT 120

Qy   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGAT 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGAT 180

Qy   181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAGCTTAGCTTAAAGACTATTAGTTTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAGCTTAGCTTAAAGACTATTAGTTTG 240

Qy   241 TTAGGTACACACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAAAA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTAGGTACACACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAAAA 300

```
ACCESSION   I17387
VERSION     I17387.1    GI:1597742
KEYWORDS    Unknown.
SOURCE      Unknown.
  ORGANISM  Unclassified.
REFERENCE   1  (bases 1 to 1305)
  AUTHORS   Saito,S., Ohkawa,S., Fujisawa,A., Iritani,Y. and Aoyama,S.
  TITLE     Poultry mycoplasma antigen, gene thereof and recombinant vectors
            containing the gene as well as vaccines utilizing the same
  JOURNAL   Patent: US 5489430-A 1 06-FEB-1996;
FEATURES             Location/Qualifiers
     source          1..1305
                     /organism="unknown"
                     /mol_type="unassigned DNA"

ORIGIN

Query Match             96.8%;  Score 1264;  DB 6;  Length 1305;
Best Local Similarity   98.3%;  Pred. No. 2.6e-174;
Matches 1284;  Conservative  0;  Mismatches  21;  Indels  1;  Gaps  1;

Qy      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA  60
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAAACACAAAATCTTCTAACAA  60

Qy     61 AATCCTAAATAAAATAAAGCCGTTAAATTAACTAAAAAATTAAAAAAATAAAAATAATGG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAAATAAAATAAAGCCGTTAAATTAACTAAAAAATTAAAAAAATAAAAATAATGG  120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCATCTTTTAGAT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCATCTTTTAAGAT  180

Qy    181 ATAAATATATCTTAATATCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATAAATATATCTTAATATCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG  240

Qy    241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTGTATGTTAACAGAT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTGTATGTTAACAGAT  300

Qy    301 GATGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGAATGGAGTTAACTGAT  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGAATGGAGTTAACTGAT  360

Qy    361 CTAATCAATGCTAAAGCCAATTACCAGCAGAGGCTTCACTACAAGATTGAAGCT  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAGCCAATTACCAGCAGAGGCTTCACTACAAGATTGAAGCT  420

Qy    421 AGTTTATCATCTGCTTATAGTGAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTTTATCATCTGCTTATAGTGAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA  480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540

Qy    541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC  600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT  660

Qy    661 CGCAATAATTTAGTCGGATCATATACAAGCAGTAGTTAATAACACTAGAT  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGTAATAATTTAGTCGGAATGCTTTAGATTCAGTAGTTAATAACACTAGAT  720

Qy    721 CCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCACTAAATGGGGGAATGCTTTAGATTCAGTAGAGATTACTACAGTTAATCGGAATATT  780
```

```
Qy    301 GATGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGAATGGAGTTAACTGAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGAATGGAGTTAACTGAT  360

Qy    361 CTAATCAATGCTAAAGCCAATTACCAGCAGAGGCTTCACTACAAGATTGAAGCT  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAGCCAATTACCAGCAGAGGCTTCACTACAAGATTAACTGAT  420

Qy    421 AGTTTATCATCTGCTTATAGTGAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTTTATCATCTGCTTATAGTGAGCTGAAACAGTTAATAACAATAACCTTAATGCAACACTA  480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540

Qy    541 TTGGTTAAAGCAGCTGATGCTGATAACGTTGGATTACAATACAAATAAATAATGGAAAT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTGGTTAAAGCAGCTGATGCTGATAACGTTGGATTACAATACAAATTAAATAATGGAAAT  600

Qy    601 GTTCAACAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACACAGCTAATCCAACT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTTCAACAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACACAGCTAATCCAACT  660

Qy    721 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCCAGGTT  1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    721 -CAGCAGTTGATGAGATTAAAGTTGCTTAAAATCGTTTTATCCAGGTT  1305
```

```
Qy    301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGCGAATGAGTTAACAGAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GACGCCAAACCCAATTAGCCAAACCCAATTACCAGCAGCGCGCGAATGAGTTAACTGAT  360

Qy    361 CTAATCAATGCTAAAGCCGATGATGAACATTAGTTCACTACAAGACTACAAGAATTGAAGCT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAGCCGATGATGAACATTAGTTCACTACAAGACTACAAGAATTAACTGAT  420

Qy    421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAATAACAATAACCTTAATGCAACACTA  480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540

Qy    541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAAACC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAACGACTTTGATAATGAACATCCAAATTAGTTGAAGCATACAAAGCACTAAAAAACC  600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCACTTCAACTGCTTATAATCCAACT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCACTTCAACTGCTTATAATCCAACT  660

Qy    661 CGCAATAATTTAGTTGCTAAGTTACTACAGTTTTTTGTTGCTTTGTGTCTTTTAGTCGTGA  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGTAATAATTTAGTTGCTAAGTTACTACAGTTTTTTGTTGCTTTGTGTCTTTTAGTCGTGA  720

Qy    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840

Qy    841 TTTATTAAAAAGTGATTCAAACTACAGTTTGCTTTTGTTGCTGTGATTGTAACACCCGTCAAT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTATTAAAAAGTGATTCAAACTACAGTTTGCTTTTGTTGCTGTGATTGTAACACCCGTCAAT  900

Qy    961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020

Qy   1021 AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACAAACGGAG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACAAACGGAG  1080

Qy   1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG  1140

Qy   1141 TTGGTTAAAGCAGCTGATGCTGATAACGTTGGATTACAATACAAATAAATAATGGAAAT  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGTTAAAGCAGCTGATGCTGATAACGTTGGATTACAATACAAATTAAATAATGGAAAT  1200

Qy   1201 GTTCAACAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACACAGCTAATCCAACT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCAACAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACACAGCTAATCCAACT  1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCCAGGTT  1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 -CAGCAGTTGATGAGATTAAAGTTGCTTAAAATCGTTTTATCCAGGTT  1305
```

```
RESULT 3
I17387
LOCUS       I17387             1305 bp    DNA        linear   PAT 06-OCT-1996
DEFINITION  Sequence 1 from patent US 5489430.
```

```
                    /organism="Mycoplasma gallisepticum"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:2096"

ORIGIN

     Query Match      96.7%;  Score 1262.4;  DB 6;  Length 1387;
     Best Local Similarity  98.2%;  Pred. No. 4.4e-174;
     Matches 1283;  Conservative  0;  Mismatches 22;  Indels 1;  Gaps 1;

Qy     1 AAAAACATCAGATTGTTAATCTGTATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AAAAACATCAGATTGTTAATCTGTATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA    60

Qy    61 AATCCTAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATC   120
         ||||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||
Db    61 AATCCTAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATAGTTTTTCTTATC   120

Qy   121 AACCAAAATTCTCAGTAATAAACGCTTATTTTATTTTTATTTTTAGTTGTCATCTTTAGAT   180
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   121 AACCAAAATTCTCAGTAATAAACGCTTATTTTATTTTTATTTTTTTACTCATCTTTTAGAT   180

Qy   181 ATAAATATATCTTAATATTCATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG    240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATAAATATATCTTAATATTCATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG    240

Qy   241 TTAGTACAACATCCTTCTTAGCATTGGGATTCAGCTGTATGTCTATTACTAATAAAAA     300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTAGTACAACATCCTTCTTAGCATTGGGATTCAGCTGTATGTCTATTACTAATAAAAA     300

Qy   301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCAGGCGGCGAAATGAGTAACAGAT   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCAGGCGGCGAAATGAGTAACTGAT   360

Qy   361 CTAATCAATGCTAAAGCGATGACATAGCTTCACTACAAGATCATGCCAAGATTGAAGCT   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTAATCAATGCTAAAGCGATGACATAGCTTCACTACAAGACTATGCTAAGATTGAAGCT   420

Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTAAAAACTAATTGAAATGCAACAATAACCTTAATGCAACATTA   480
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   421 AGTTTATCATCTGCTTATAGTGAAGCTAAAAACAGTTAACAATAACCTTAATGCAACACTA   480

Qy   481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAGCTAATACGAT   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAGCTAATACGAT   540

Qy   541 AAAACGACTTTGATAATGAACACCAAATTAGTTGAAGCATACAAAGCACTAAAAACC   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAAACGACTTTGATAATGAACATCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC   600

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT   660
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCAACTGCTTATAATCAGATT   660

Qy   661 CGCAATAAATTTAGTGGATCATACAAATAAAGCTAGTAGTTTAATAACTAAAAACACTAGAT   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGTAATAAATTTAGTGGATCATACAAATAAAGCTAGTAGTTTAATAACTAAAAACACTAGAT   720

Qy   721 CCACTAAATGGGGAACGCTTTAGATTCTAATGATACTACAGTTAATCGGAATATT   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCACTAAATGGGGAATGCTTTAGATTCTAATGATACTACAGTTAATCGGAATATT   780
```

```
Qy   781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT   840

Qy   841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTGCTTTTAGTGTAGGGACTTTTACAAACGCT   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTGCTTTTAGTGTAGGGACTTTTACAAACGCT   900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAAT   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAAT   960

Qy   961 TATAAAATATGCAAGAGGACCGTTTGGAGTGTGATGAGAACCTTCAAGTAGAATTCTTGCA   1020
         |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db   961 TATAAAATATGCAAGAGGACCGTTTNNNAATGGTGATGAACCTTCAAGTAGAATTCTTGCA   1020

Qy  1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTAGCTGGAACAAACACGAAG   1080
         |||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db  1021 AACACGAATAGTATCACAGATGTTTCTNNNATTTATAGTTAGCTGGAACAAACACGAAG   1080

Qy  1081 TACCAATTTAGTTTTAGCAGCTGATGCTAATAACGTTCGATTACAACTGGTTATTTATATTTCCCTTATAAG   1140
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  1081 TACCAATTTAGTTTTAGCAGCTGATGCTAATAACGTTGGATTACAACTGGTTATTTATATTTCCCTTATAAG   1140

Qy  1141 TTGGTTAAAGCAGCTGAGTTGCCACTTCAACTAGTGCAAATAATAACTACACAGCTAATCCAACT   1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TTGGTTAAAGCAGCTGAGTTGCCACTTCAACTAGTGCAAATAATAATAACTACACAGCTAATCCAACT   1200

Qy  1201 GTTCAACAAGTTGAGTTGCCACTTCAACTTCAACTAGTGCAAATAATAATATACTACAGCTAATCCAACT   1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTTCAACAAGTTGAGTTGCCACTTCAACTTCAACTAGTGCAAATAATAATATACTACAGCTAATCCAACT   1260

Qy  1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTTTATCAGGTT   1306
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTTTATCAGGTT   1305
```

```
RESULT 4
E09301
LOCUS       E09301                  1387 bp    DNA        linear   PAT 29-SEP-1997
DEFINITION  DNA encoding Mycoplasma TM-16.
ACCESSION   E09301
VERSION     E09301.1  GI:22025928
KEYWORDS    JP 1995133295-A/2.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 1387)
  AUTHORS   Mori,H., Saito,S., Okawa,S., Funato,H., Iritani,K., Aoyama,S. and
            Takahashi,K.
  TITLE     NEW ANTIGEN PROTEIN, ITS GENE, RECOMBINANT BACULOVIRUS AND ITS USE
  JOURNAL   Patent: JP 1995133295-A 2 23-MAY-1995;
            NIPPON ZEON CO LTD, SHIONOGI & CO LTD
COMMENT     OS   Mycoplasma gallisepticum
            PN   JP 1995133295-A/2
            PD   23-MAY-1995
            PF   27-AUG-1993 JP 1993213102
            PI   MORI HAJIME, SAITO SHUJI, OKAWA SETSUKO, FUNATO HIRONO, PI
            IRITANI KOICHI,
            PI   AOYAMA SHIGEMI, TAKAHASHI KIYOTO
            PC   C07K14/30,A61K39/00,A61K39/00,C12N7/01,C12N15/31,C12P21/02, PC
            (C12P21/02,
            PC   C12R1:92);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1..1387
            FT                   /organism='Mycoplasma gallisepticum'
            FT   source          202..1308
            FT                   /product='TM-16'.
FEATURES             Location/Qualifiers
     source          1..1387
```

```
Qy   961  TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020

Qy  1021  AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACAAACACGAAG  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACAAACACGAAG  1080

Qy  1081  TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTCCCTTATAAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTCCCTTATAAG  1140

Qy  1141  TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAATAATGGAAAT  1200
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  1141  TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAATAATGGAAAT  1200

Qy  1201  GTTCAACAAGTTGAGTTGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GTTCAACAAGTTGAGTTGTTGCCACTTCAACTAGTGCAAATAATAATAATCCAACT  1260

Qy  1261  CCAGCAGTTGATGAGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT  1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 -CAGCAGTTGATGAGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT  1305
```

```
RESULT 5
AE016968
LOCUS       AE016968             302070 bp   DNA     linear   BCT 03-SEP-2003
DEFINITION  Mycoplasma gallisepticum strain R section 2 of 4 of the complete
            genome.
ACCESSION   AE016968 AE015450
VERSION     AE016968.1  GI:31541264
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum R
  ORGANISM  Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 302070)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE     The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
  PUBMED    12949158
REFERENCE   2  (bases 1 to 302070)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzell,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathobiology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eaglevile Road U-89, Storrs, CT
            06269-3089, USA
FEATURES             Location/Qualifiers
     source          1..302070
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:233150"
     gene            79..3588
                     /locus_tag="MGA_1011"
                     /note="MGA_1011; MGR_224"
     CDS             79..3588
                     /locus_tag="MGA_1011"
                     /codon_start=1
                     /transl_table=4
                     /product="conserved hypothetical"
                     /protein_id="AAP56566.1"
                     /db_xref="GI:31541265"
                     /translation="MMNQNKKNNLPKKPLVKNKGIDLDDFLSEEIAVLDNFHLTKPKK
                     QQPLLLENKRAKKVSQPVVVKPGEKKVKAKKAKVVITDPRILKKRSRLLTIGFSLIGF
                     GITGPLLSFAILNQIDKQTDQSAPITRPIDSPTDTLQSSPQFTHDNSLPAFYNSYQL
                     NKQSEINGIDLSNQRTLHDLPIASFMHHNKLVRLLNQQLHFNYSWLYSDIAYLIRKTY
                     DENKQAFIPLKTTDLQTKSSDNNPLHVYLTANFEFRNQTDSPQTVFVRYNQRYEBKQ
```

```
gene            3664..5019
                /locus_tag="MGA_1014"
CDS             3664..5019
                /locus_tag="MGA_1014"
                /note="MGA_1014; MGR_225"
                /codon_start=1
                /transl_table=4
                /product="unique hypothetical"
                /protein_id="AAP56567.1"
                /db_xref="GI:31541266"
                /translation="MTPYQKEIPDYDDEGFVDSIFHDDETDRFDEDQQDHFGFDSRTN
                RIKKNVNYDQHDLQTLDVSRNKNRFPDKLVRSTKTYNFIAIWLNLIPLITFVLAIVFF
                YLNKFSPQQSESTVVSVNLFLTIALAAVSYFISVILYGFVKMHNKQIKYSLTEKSLDY
                YSIFKLYVSYFLLGWIPLLGVIISGHIALQNLDDSDVSDNMLWYQESISDLLIDM
                NGYENSYRELESKKVYLDQELQRLNYEKQKLNAKGNBLAVVSSNALPNFKNSKQLLNY
                VSNSGFLSASLVEQVKSAITQLREBRRKLRRENRKLRALLSKANNIDEQPLLTSSNN
                LFNDGEQESLFNLTDQPSRKNEGLNLENPKDRIDWISENLNNDRAIVEIGQIKKETAQ
                NLRSRAPKVVKKNNQPTYKVMVNNKKEAKVGRNQEBSDLTKQSNLLNFVNARIWVNQ
                L"
tRNA            5196..5269
                /product="tRNA-Trp"
                /note="codon recognized: UGG; tRNA genes predicted using
                FASTRNA software of the Institut Pasteur, France;
                functional prediction of the tRNAs is performed according
                to Ushida C. and Muto"
gene            6561..7661
                /gene="hatA"
                /locus_tag="MGA_1017"
CDS             6561..7661
                /gene="hatA"
                /locus_tag="MGA_1017"
                /note="MGA_1017; high affinity transport protein A (HatA)
                [Mycoplasma gallisepticum]; MGR_226"
                /codon_start=1
                /transl_table=4
                /product="HatA"
                /protein_id="AAP56568.1"
                /db_xref="GI:31541267"
                /translation="MSKTIKKLFLISSSFLLSSAFVSSCAKANDANNQGTIQLKVDLK
                IDHDANAIINDYLKNLSQTVSQKLNKPVSITSNFTBDGNTSIDRVKSNLTDLAFVSSS
                AIKNDTDYITKIQTLTRSFKFDNTPDFYEAGTLDQKARKISELFNEIPYRNWTDENR
                MWDGNKYQFFYGPEDDLISFYRGMILIGGDQDTLAKIKEAWAQKNWATFSSYGIGGGR
                TSSNGRYIYPLNLLKFHFNDPNLTISDYQSVRGREIGTNFKVHIVFDDMNSFAWTQNK
                NKNVPNYTNFPNDNNTKVEILSLTDPALNIGVFNKRLTNDVINAVSEBSIVELAKKNQ
                DTYGPTVGYNGYRVIKDVNKEVYQFINKANIN"
gene            7661..8407
                /gene="hatB"
                /locus_tag="MGA_1018"
CDS             7661..8407
                /gene="hatB"
                /locus_tag="MGA_1018"
                /note="MGA_1018; high affinity transport protein B (HatB)
                [Mycoplasma gallisepticum]; MGR_227"
                /codon_start=1
                /transl_table=4
                /product="HatB"
                /protein_id="AAP56569.1"
```

```
IQPKGILNLKLSLSASNDIDPLKDPDYIGSIDPVDLAPGYGYINAVVTKDLYHNYYL
NWNLSLBLSMNDSNYRFSNFIFNNNVPSMNILLKKTIPGKDPYQAANDLVGENKLLN
RDLTPEQISRNLKFYFSNEFHSQIEWFNLIRDVFKWLIVNNTNKKLYLVLSDNAAVF
AKLFSYNFIFKLNINYANSKNAILKLITDALSSNPTIVNAKNVYEVLYDNFPDAFKLLIN
TYFFDVSAYSNLIDSLTKDVNSAFDPYQRLYELMPILRQLISNFPTKIKYFBIIRIL
GYRDDAGNLIPIPRFNTYLINTIFESKRLLNDIYDQLLNISNVDFDMATANTNNGSES
WSRFTRLFNFLFIQNRSTLVNLFTQSTNYFDDLITLFISNNFNFLFQLLKLFGINVNT
LDLFFKTALDLIFLKNPEILSRDNLKAKIINLFNWIIKSLDQTNINKIKFGLYKPENF
DNLVLTQSPTTLKVKKLDLGFKLDINGLQIPYGVLSALFEFLPNIKIVDPFKYIVDDK
RYEBLKEKGYQQAINEGYSSLLNAGGIFSFKWAGDPAKNFPSSVFDDIWKNRILKGVS
DNNRTIKDLANEIFSGTNNNDPNRVIATINGAGYFKWSGENVDILPFFEILPSNINQA
QIVYQLYNVDQBKDFSQIYQNLKLHFPRYSYRQPGIPPSDLSKKIYQSISGMVDQTLN
YSLNRKYLTKTRIILPGGDGKKNHLILQENNPYENVALFTYNANQSLINQYDRMNBQD
KRKLREQIMITDRRARVKGIDRDPKTRETSLLPQSLRMLDSLIQIKGVNPSLVIHTN
WSFSTISERISLNFITBVKPGLISKQVSINLAYETFLSAFSYQVLLPYPIIDLSNPNN
PVFKLKBEISHLVFSLBLPGAGB"
```

```
                  /db_xref="GI:31541268"
                  /translation="MSNTNHIKFLDVSIKHKRYKGFNVDHINPELKGSEILMLIGKSG
                  SGKTTLLNSITDKLVSQGEIVFNDQNLPSLSKHKIRKIAKPIGFLDQIPNLILDDYV
                  YENIIRSIRKPWYYKFNITPISVLNKVBEVLKTLDLSDHINKLIGKLSGGQXQRVDI
                  AKIFFQKPKLLLIDEPTTGLDYLNSEIVIREIKKLSEQFKAPVIISIHDLEIVKKFAD
                  KILIIQNKKPILINNSADLNIDEIRKKYET"

gene              8379..9998
                  /gene="hatC"
                  /locus_tag="MGA_1019"

CDS               8379..9998
                  /gene="hatC"
                  /locus_tag="MGA_1019"
                  /note="MGA_1019; high affinity transport protein C (HatC)
                  [Mycoplasma gallisepticum]; an alternative start codon may
                  be 27 nucleotides upstream; 'tta aca att ctg cgg atc tta
                  ata tcg ATG' / 'LTILRILISM'; MGR_228"
                  /codon_start=1
                  /transl_table=4
                  /product="HatC"
                  /protein_id="AAP56570.1"
                  /db_xref="GI:31541269"
                  /translation="MRLEKNMRHNKFFQLLKINKFTVFFTVFLCIISAIVIKLGIN
                  AHGLAVFKNNISNFFNQSAVNKSLLTKSFEQIWITIKYTATGTFIGFLGFLLGYLS
                  SLRITNNFTAIFVKLGIIFPRSFPVVIFNLPNTSFNSDLSAIIIISWFSMLWNAKYI
                  ADYIENSKVEQFKKFVNRSQNHLSSFINNIFIDIKSKIIVLFAYSLESNFRWTTILSA
                  VGLIGIGQLINDPISINNDFASTLIPLLVMVFLLINEGALYLFETYLVARKSYPKTK
                  NLNKLLTIKKYLVLVLIFILIITFSIYSIFSIRLRTIIIITLGAYGTSVIGQLLDYIYRGSIENL
                  TTFNENPLLMIILLTLQAILLGIVFIFSLLFATLCSNLLNRYVSLFFKALFLIRTI
                  PLIIVFRLFNPLFNSGISTIVFIGSLYFSTSLAKKIYVLINSINWTVVTSLKSKLYTN
                  FEILRVYIIPSIRKDLVTVYQLEFSILRTIIITLGAYGTSVIGQLLDYIYRGSIENL
                  GSYVISIMVYFQVIDLLSIMVRYKKFFYTNK"

gene              10143..11567
                  /gene="proS"
                  /locus_tag="MGA_1022"

CDS               10143..11567
                  /gene="proS"
                  /locus_tag="MGA_1022"
                  /note="MGA_1022; [J] COG0442 Prolyl-tRNA synthetase;
                  MGR_229"
                  /codon_start=1
                  /transl_table=4
                  /product="ProS"
                  /protein_id="AAP56571.1"
                  /db_xref="GI:31541270"
                  /translation="MAFKKNLASSKDDNISDWYTEVISKGNLIHYSNIKGFMSFLPNG
                  WKIWQLIKKELTKEFDQRDILELCLPSLISQNDFPLLEKKHLEGPAPELFIVTKISNET
                  EKYILRPTSEIAFCNLWRETLRNYRDLPIKHNQWTSVFRVEKNTRPFLRNSERFWHEI
                  HSCFETEBEQSNDFAVDIWKLYQRFIKDLLCIPLVAGEKTEAEHKFAGAKTTYTVETIMP
                  DGQALQSATSHNLSQNFSKAFDIRYQTKNNDYQNVFSMSAGVSTRIIGAIIMTHGDDD
                  GLVFPTKVAPYHISLNCIFDDTNQELNAKLKELANKYSQKYRVHLNVNKDSTGEIIKN
                  SQLRGDCCVLLMGPNDLKKNEIVPIDRITKQKQFINLDHLDQKLBELFSTFDQKLYQK
                  AKAVFETKVDFAQTFBEFEQKIASGKFVRVFYCNEDLYEKQIKEKTGASSRCIIKYLD
                  EQTQERCFISNKKAKVEIYFARSY"

gene              11768..12190
                  /locus_tag="MGA_1027"

CDS               11768..12190
                  /locus_tag="MGA_1027; MGR_230"
                  /codon_start=1
                  /transl_table=4
                  /product="unique hypothetical"
                  /protein_id="AAP56572.1"
                  /db_xref="GI:31541271"
                  /translation="MKMFKVKIAKKIFHTFVKKQNHILNRQPERSISMNRKIINDAKKY
                  LERYDPYRKRIENRQIIKSVFNHDIGVDNQSEMQSYVDLIDATLKRLKKDEKDILIDV
                  FLKRKHRNEFNYSQASPYRLLNKACCSLFYSLGVVIPN"

gene              12194..13276
                  /locus_tag="MGA_1029"

CDS               12194..13276
                  /locus_tag="MGA_1029; MGR_231"
                  /codon_start=1
                  /transl_table=4
```

```
                  /product="unique hypothetical"
                  /protein_id="AAP56573.1"
                  /db_xref="GI:31541272"
                  /translation="MLISYNIDSNNNSKPIVINKGWWWRLNPLRNKDGTIVVDKNSN
                  PIYFRALLSFQNTWSDQNNLNLYFEILFENNSILPPNTIQSNVFIAGNVPSVFSEIT
                  NEETDGYQKAQIKVSVPLBKLTKERFBEGIDISVVIDLYLQNDILESGIRDSLESNALR
                  LLLENLKDARLVTKIKLNNFLFPRDRVGNRDDADYQVQEKKAISWYQPNYSFLSNLD
                  QIRDLLTINVDAVTLFKSRVVFDYSTTNGSTKSLVIENTNQITYPPNTDIKLALDT"
```

```
Query Match            90.3%;     Score 1179.4;  DB 1;   Length 302070;
Best Local Similarity  95.3%;     Pred. No. 8.4e-163;
Matches 1249;   Conservative   0;   Mismatches   56;   Indels   6;   Gaps   3;

Qy      1 AAAAACATCAGATGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA     60
          ||||| |||||||||||||||||||||||||| |||||| |||||||||||| |||||
Db 189913 AGAACAATCAGATGTTAATCTGATATCTTGCTT-AAAAACACAAAATCTTTTAACAA 189971

Qy     61 AATCCTAAATAAATAAAGCCGTTAAATAACTAAAAAAATTAAAAAAAATGGTTTTTCTTATC    120
          |||||||||||||||||||||||||||||||||||  |||||||  ||||||||||||||||
Db 189972 AATCCTAAATAAATAAAGCCGTTAAATAACTAAAAAAATTATTTAATTAAAAATGGTTTTTCTTATC 190031

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTAT-TTTTATTTTAGTCATCTTTAAGA    179
          ||||||||||||||||||||||||||||||||||  ||||||||| ||||| ||||||
Db 190032 AACCAAAATCTTTAGTAATAAACGCTTATTTATTTATTTTTATTTTGCTCAATCATTTAAGA 190091

Qy    180 TA-----TAAATATATCTATATCTCATGAATAAGACAAAAGAATCATCTTAAAGACTATTA    235
          ||      ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 190092 TATATCTATATATATATCTAATATTCTATGAGTAAAAAAAGAATCATCTTAAAGACTATTA 190151

Qy    236 GTTTGTTAGGTACACACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTA    295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190152 GTTTGTTAGGTACACACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTA 190211

Qy    296 AAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTGAAGCAGCGCGAATGGAGTTAA    355
          |||||||| |||||||||||||| |||| |||||||||||||||||| |||| |||||||
Db 190212 AAAAAGACGCAAATCCAAATCAAATAATAGCCAAACCCAATTACAAACAGCGCGCGAATGGAGTTAA 190271

Qy    356 CAGATCTAATCAATGCTCAAAGCAATGACATTAGCTTCACTACAGACATATGCCAAGATTG    415
          | |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db 190272 CTGATCTAATCAATCAAAGCAATGACATTAGCTTCACTACAGACAAGAATATGCCAAGATTG 190331

Qy    416 AAGCTAGTTTATCATCTGCTTATAGTGTGAAGCTAAATCAATTAGAATCAGCCATCAACGCAA    475
          |||||||||||||||||||||||||||||||||||||||||| ||||||||| ||||||||
Db 190332 AAGCTAGTTTATCATCTGCTTATAGTGTGAAGCTGAAACAGTAACAATCAATTAGAATCAGCCATCAACGCAA 190391

Qy    476 CATTAGAACAACTAAAAATGGCTAAAACTAATTAGATGAATCAGCAGCCATCAAAGACACTAA    535
          ||||||| |||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db 190392 CACTAGAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCAGCCATCAACCACCAAGCTAATA 190451

Qy    536 CGGATAAAACAGACTTTGATAATGAACACCCAAAATTAGTTGAAGCATACAAAGCACTAA    595
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 190452 CGGATAAAACCACTTTGATAATGAACACCCAAAATTAGTTGAAGCATACAAAACACTAA 190511

Qy    596 AAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATC    655
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db 190512 AAACCACTTTATCACAACGTGCTACTAACCTTGAAGGTTTAGCATCAACTGCTTATATATC 190571

Qy    656 AAATTCGCAATAATTTAGTGGATCTATACAATAAGCTAGTAGTTAAATAACTAAAAACAC    715
          |||||| ||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db 190572 AGATTCGTAATAATTTAGTGGATCTATACAATAAGCTAGTAGTTAATAACTAAAAAACAC 190631

Qy    716 TAGATCCACTAAATGGGGAACGCTTTAGATTCAATGAGATTACTACAGTTAATCGGA    775
          |||||||||||||||||||||||||| ||||| |||||||||||||||||||||||||
Db 190632 TAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAGA 190691

Qy    776 ATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTA    835
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||| ||||
Db 190692 ATATTAATAATACGTTATCAACTATTAATTGAACAAAAGACTAATGCTAATGCATTAGCTA 190751

Qy    836 ATAGTTTATTAAAAAGTGATTCAAAATAATGAACATAATGAACAAAGTTTGTGGGGACTTTTACAA    895
          ||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 190752 ATAGTTTTATCAAAGAAGTGATTCAAAATAAATAAACAAAGTTTGTAGGAATGTTTTACAA 190811
```

```
Qy    241 CTTCACTACAAGACTATGCCAAGATTGAAGATTGAAGAGCTAGTTTATCATCTGCTTATATGTAAGCTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190812 ACACTAATGTTCAACCTTCAAACTATAGTTTGTTGTCTTTAGTGCTGATGTAACTAATT 508

Qy    301 AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACATTAGAACAACTAAAAATGGCTAAAACTAATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    956 TCAATTATAAATATGCAAGAAGGACCGTTTGCAATGTGATGAACCTTCAAGTAGAATTC 1015

Qy    361 TAGAATCAGCCATCAACAAGCTAATACGCTAATACGAATAAAAACGACTTTGATAATGAACACCCAA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190872 TTAATTTATAAATATGCAAGAAGGACCGTTTGAAATGTGATGAACCTTCAAGTAGAATTC 190931

Qy    421 ATTTAGTTGAAGCATACAAAGCACTAAGACACTAAGACACCACTTAGAACACACGTGCTACTAACCTTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1016 TTGCAAACACGAATAGTATCACAGATGTTTCTGATGTTTCTGATTTATAGTTAGCTGGAACAAACA 1075

Qy    481 AAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAAATTCGGCAATAATTTAGTGGATCTATACAATA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190932 TTGCAAACACGAATAGTATCACAGATGTTTCTTGATTTATAGTTTAGCTGGAACAAACA 190991

Qy    541 AAGCTAGTAGTTTAATAACTACTAAAACTACTAGATCCACTAGATCCACTCAAATAATAAACTAAACTAAATAATG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1076 CGAAGTACCAAATTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATTATTTCCCTT 1135

Qy    601 CTAATGAGATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190992 CGAAGTACCAATTAGTTTTAGCAACTACGCGTCCATCAACTGGTTATTTATTATTTCCCTT 191051

Qy    661 AAAAGAACTAATGCTGATGCATTACCAGCTAATAGCATTATGTTTTATTAATAAGAAGTGAATTCAAAATAATA 720
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1136 ATAAGTTGGTAAAGCAGCTAATAACGTTGGATTACAATACAATTAAATAAATG 1195

Qy    721 AACAAAGTTTTGTAGGAGTTTACAAACGCTAATGTTCAACCTTCAACCTTCAAACTACAGTTTTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 191052 ATAAGTTGGTAAAGCAGCTAATGTGCTAGTAATGTTCGATTACAATACAAACTAAATAATG 191111

Qy    781 TTGCTTTTAGTGCTGATGTAACACCTGTTAATTATAATAAAATATGCAAGAAGACGGTTTGAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1196 GAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTACACAGCTAATC 1255

Qy    841 ATGGTGATGAGAACCTTCAAGTAGAATTCTTCGCAAACAGCAATAGTCACGAATGTTCTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191112 GAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTCAAATAATAAACTAAACTACAGCTAATC 191171

Qy    901 GGATTTATAGTTTAGTTATCTGGAACAAACACCATTGGACAATAAATAAATGGAATGTTTAGCAACTACGGTC 960
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1256 CAACTTCCAGCAGTTGATGAGATTAAAATCGTTTTATCAGGTT 1306

Qy    961 CATCAACTGGTTTATTTATTTTCCCTTATATATTTCCCTTATATATTCCCTTATAAGTTAATG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 191172 CAACTTCCAGCAGTTGATGAGATTAAAAATCGTTTTATCAGGTT 191222
```

RESULT 6
AR035278

```
LOCUS       AR035278     2014 bp    DNA      linear    PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5871742.
ACCESSION   AR035278
VERSION     AR035278.1  GI:5951946
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2014)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 7 16-FEB-1999;
FEATURES    Location/Qualifiers
  source    1..2014
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match       83.2%;   Score 1086.8;  DB 6;  Length 2014;
Best Local Similarity  96.7%;   Pred. No. 1.2e-148;
Matches 1120;  Conservative 0;  Mismatches 37;  Indels 1;  Gaps 1;
```

```
Qy    150 TTTATTTTATTTTGT-CATCTTTAAGATATAAAATATCTTAAATATTCTAATATTCTATGAATA 208
          |||||||||||||||  |||||||||||||||||||||
Db      1 TTTATTTTATTTTGTAAATCTTTAAAATCTTTAAAATATAATATTTAAATATTCTAATATTCTATGAATA 60

Qy    209 AGAAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTCTTAGCATTG 268
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAAAAAGAATCATCTTAAAGACTATTAGCTTGTTAGGTACAACATCCTTTCTTCTTAGTATTG 120

Qy    269 GGATTTCTAGCTGCTGTATGTGTATGTCTATTACTAATACTACTAAAAAGATGCAAAATGGCCAAACCC 328
          |||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGATTTCTAGCTGCTGTATGTGTATTACTAATACTACTAAAAAAGAATGCAAAACCCAAATAATGCCAAACCC 180

Qy    329 AATTAGAAGCAGCGCGAATGAGTGAACTAACAGATCTAATCAATGCTAAAGCGATGACATTAG 388
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATTAGAAGCAGCGCGAATGAGTGAACTAACAGATCTAATCAATGCTAAAGCGATGACATTAG 240

Qy    389 CTTCACTACAAGACTATGCCAAGATTGAAGATTGAAGAGCTAGTTTATCATCTGCTTATATGTAAGCTG 448
          ||||||||||||||||||||||||||||||
Db  190812
```

RESULT 7
AX665164

```
LOCUS       AX665164    1152 bp    DNA    linear    PAT 26-MAR-2003
DEFINITION  Sequence 1 from Patent EP1275716.
ACCESSION   AX665164
VERSION     AX665164.1  GI:29290294
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
```

```
Qy   1049 GGATTTATAGTTTAGCTGGAACAAACACAGAAGTACCAATTTAGTTTTAGCAACTATGGTC 1108
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAATTTATAGTTTATCTGGAACAAACACAGAAATACCAATTTAGTTTTAGCAACTACACGGTC 960

Qy   1109 CATCAACTGGTTTATTTATTTTCCCTTATATATTTCCCTTATATATTCCCTTATAATAAACG 1168
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CATCAACTGGTTTATTTATTTATTTTCCCTTATATATTTCCCTTATATATTCCCTTATAAATG 1020

Qy   1169 TTGGATTACAATACAAATAAAATAAAATAAAACTAAATAATGGAAATGTTCAACAAGTTGCCACTTCAA 1228
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTGGATTACAATACAAATAAAACTAAATAATGGAAATGTTCAACCAGTTGAGTTTGCCACTTCAA 1080

Qy   1229 CTAGTGCAAATAATAATACTACAGCTAATCAATGCTAAAGGATTAAAAGTTGCTA 1288
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTAGCGCAAATAATAATACTACAGCTAATCAATGCTAAAGGATTAAAAGTTGCTA 1140

Qy   1289 AAATCGTTTTATCAGGTT 1306
          ||||||||||||||||||
Db   1141 AAATCGTTTTATCAGGTT 1158
```

```
REFERENCE    1
AUTHORS      Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki.Y.
TITLE        Modified dna molecule, recombinant containing the same thing, and
             uses thereof
JOURNAL      Patent: EP 1275716-A 1 15-JAN-2003;
             Zeon Corporation (JP)
FEATURES             Location/Qualifiers
source       1..1152
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:2096"
                     /note="TM-1 gene (after EcoRI)"
ORIGIN

Query Match      78.7%;  Score 1028;  DB 6;  Length 1152;
Best Local Similarity  100.0%;  Pred. No. 4.8e-140;
Matches 1028;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    279 CTGTATGTCTAATTACTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGC 338
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6 CTGTATGTCTAATTACTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGC 65

Qy    339 AGGCGCAATGGAGTTAACAGATCTAACAGATCAATGCTAAAGCGATGACATTAGTTCACTACA 398
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 AGCCGCAATGGAGTTAACAGATCTAACAGATCAATGCTAAAGCGATGACATTAGTTCACTACA 125

Qy    399 AGACTATGCCAAGATTGAAGCTAGTTTATCATCGTGCTTATAGTGAAGCTGAAACAGTTAA 458
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 AGACTATGCCAAGATTGAAGCTAGTTTATCATCGTGCTTATAGTGAAGCTGAAACAGTTAA 185

Qy    459 CAATAACCTTAATGCAACATTAGAACACAATTAGAAATGGCTAAAACTAATTTAGAATCAGC 518
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 CAATAACCTTAATGCAACATTAGAACACAACTAGAAATGGCTAAAACTAATTTAGAATCAGC 245

Qy    519 CATCAACCAAGCTAATACGGATAAAAACGACTTTGATAAATGAACACCCAAATTTAGTTGA 578
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 CATCAACCAAGCTAATACGGATAAAAACGACTTTGATAAATGAACACCCAAATTTAGTTGA 305

Qy    579 AGCATACAAAGCACTAAAAACCACTTTAGAACACACGTGCTACTAACCTTGAAGGTTTGTC 638
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 AGCATACAAAGCACTAAAAACCACTTTAGAACACACGTGCTACTAACCTTGAAGGTTTGTC 365

Qy    639 ATCAACTGCTTATAATCAAATTGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAG 698
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 ATCAACTGCTTATAATCAAATTGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAG 425

Qy    699 TTTAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCTTTAGATTCTAATGATGAT 758
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 TTTAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCTTTAGATTCTAATGATGAT 485

Qy    759 TACTACAGTTAATCGGAATATTAATAAATACGTTATCAACTATTAATGAACAAAGACTAA 818
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 TACTACAGTTAATCGGAATATTAATAAATACGTTATCAACTATTAATGAACAAAGACTAA 545

Qy    819 TGCTGATGCATTATTCTAATAGTTTTATTATTGAAAAGTGATTCAAAATAATGAACAAAGTTT 878
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 TGCTGATGCATTATTCTAATAGTTTTATTATTGAAAAGTGATTCAAAATAATGAACAAAGTTT 605

Qy    879 TGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTTTGTTGCTTTTTAG 938
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 TGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAACTACAGTTTTTGTTGCTTTTTAG 665

Qy    939 TGCTGATGTAACACCCGTCAATTATAAATGCAAGAAGGACCGTTGGAATGGTGATGA 998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 TGCTGATGTAACACCCGTCAATTATAAATGCAAGAAGGACCGTTGGAATGGTGATGA 725

Qy    999 AACTTCAAGTAGAATTCTTGCAAACACGAATATCACAGATGTTTCTTGGATTTATAG 1058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 AACTTCAAGTAGAATTCTTGCAAACACGAATATCACAGATGTTTCTTGGATTTATAG 785

Qy   1059 TTTAGCTGAACAAACACGAAGTACCAATTAGTTTAGCAACTATGGTCCATCAACTGG 1118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 TTTAGCTGAACAAACACGAAGTACCAATTAGTTTAGCAACTATGGTCCATCAACTGG 845
```

```
Qy   1119 TTATTTATATTTCCCTATAAGTTGGTTAAAGCAGCAGCTGATGCTAATAACGTTGGATTACA 1178
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    846 TTATTTATATTTCCCTATAAGTTGGTTAAAGCAGCAGCTGATGCTAATAACGTTGGATTACA 905

Qy   1179 ATACAAATTAAATAATGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAA 1238
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 ATACAAATTAAATAATGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAA 965

Qy   1239 TAATACTACAGCTAATCCAACTTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTT 1298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 TAATACTACAGCTAATCCAACTTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTT 1025

Qy   1299 ATCAGGTT 1306
          ||||||||
Db   1026 ATCAGGTT 1033
```

```
RESULT 8
E09896
LOCUS        E09896                   1015 bp    DNA       linear     PAT 29-SEP-1997
DEFINITION   Mycoplasma galliseptticum TM-1 gene.
ACCESSION    E09896
VERSION      E09896.1  GI:2202065624
KEYWORDS     JP 195236498-A/2.
SOURCE       Mycoplasma galliseptticum
  ORGANISM   Mycoplasma galliseptticum
             Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE    1  (bases 1 to 1015)
  AUTHORS    Saito,S., Kyo,T., Okawa,S. and Iritani,K.
  TITLE      DETECTION OF MYCOPLASMA GALLISEPTICUM AND NEW DNA TO BE USED
  JOURNAL    Patent: JP 195236498-A 2 12-SEP-1995;
             NIPPON ZEON CO LTD, SHIONOGI & CO LTD
  OS         Mycoplasma galliseptticum
  PN         JP 195236498-A/2
  PD         12-SEP-1995
  PF         25-FEB-1994 JP 194052764
  PI         SAITO SHUJI, KYO TSUGUO, OKAWA SETSUKO, IRITANI KOICHI PC
             C12Q1/68,C12N15/09,(C12Q1/68,C12R1:35),(C12N15/09,C12R1:35); CC
             strandedness: Double;
  CC         topology: Linear;
COMMENT
FH         Key             Location/Qualifiers
FH
FT         source          1..1015
FT                         /organism="Mycoplasma galliseptticum"
FT
FT                         202..987                                               CDS
FT         misc_feature    718..741
FT                         /product='TM-1 gene product'
FT                         /note="this region is favorble for            FT
             hybridization probe to
FT                                                                  detect TM-1 gene'.
FEATURES             Location/Qualifiers
source       1..1015
                     /organism="Mycoplasma galliseptticum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2096"
ORIGIN

Query Match      75.8%;  Score 989.4;  DB 6;  Length 1015;
Best Local Similarity  98.4%;  Pred. No. 2e-134;
Matches 999;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 AAAAACATCAGATTGTAATCTGATATCTTTGCTAAAAAACACAAAATCTTCTAACAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTAAAAAAAACACAAAATCTTCTAACAA 60

Qy     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAATTAAAAAAAATGGTTTTTTCTTATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAATTAAAAAAAATGGTTTTTTCTTATC 120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTAAGAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTAAGAT 180
```

```
JOURNAL      Mycoplasma gallisepticum as a possible protective antigen
             Vaccine 11 (10), 1061-1066 (1993)
MEDLINE      94025893
PUBMED       8212828
REMARK       GenBank staff at the National Library of Medicine created this
             entry [NCBI gibbsq 138145] from the original journal article.
             This sequence comes from Fig. 2.
FEATURES             Location/Qualifiers
     source          1..1015
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2096"
     gene            202..>1015
                     /gene="TM-1"
     CDS             202..>1015
                     /gene="TM-1"
                     /note="29 kda polypeptide; This sequence comes from Fig.
                     2; conceptual translation presented here differs from
                     translation in publication"
                     /codon_start=1
                     /transl_table=4
                     /product="TM-1"
                     /protein_id="AAB28343.2"
                     /db_xref="GI:7249262"
                     /translation="MNKKRIILKTISLLGTTSFLSIGISSCMSITKDANPNNGQTQL
                     QAARMELTDLINAKARTLASLQDYAKIEASLSSAYSEAETVNNNLNATLEQLKMAKTN
                     LESAINQANTDKTTFDNEHPNIVEAYKALKTLEQRATNLEGLASTAYNQIRNNLVDL
                     YNNASSLITKTLDPLNGGMLDSNETTVNRNINNLSTINEQKTNADALSNSFIKKV
                     IQNNEQSFVGFTFNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDBPSSRI"

ORIGIN

             Query Match      75.5%;   Score 986.2;  DB 1;  Length 1015;
             Best Local Similarity 98.2%;  Pred. No. 5.9e-134;
             Matches 997;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

Qy          1 AAAAACATCAGATTGTAATCTGATATCTTGCTTAAAAAACACAAAATCTTCTAACAA    60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAACATCAGATTGTAATCTGATATCTTGCTTAAAAAACACAAAATCTTCTAACAA    60

Qy         61 AATCCTAAATAAATAAAGCCGTTAAATCTAACTAAAAAAATTAAAAAAATGGTTTTTCTTATC   120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATCCTAAATAAATAAAGCCGTTAAATCTAACTAAAAAAATTAAAAAAATGGTTTTTCTTATC   120

Qy        121 AACCAAAATCTCTAGTAATAAACGCTATTATTTATTTTAGTCATCTTTTAAGAT   180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACCAAAATCTCTAGTAATAAACGCTATTATTTATTTTAGTCATCTTTTAAGAT   180

Qy        181 ATAAATATATCTATCTATGAATAGACTATTAGTTG   240
              |||||||||||||||||||||||||||||||||||
Db        181 ATAAATATATCTATCTATGAATAGACTATTAGTTG   240

Qy        241 TTAGGTACAACATCCTTCTTAGCATTGGAATTC   300
              |||||||||||||||||||||||||||||||||
Db        241 TTAGGTACAACATCCTTCTTAGCATTGGAATTC   300

Qy        301 GATGCAAACCCAAATAATGGCCAAACCCAATTACAAGAAT   360
              ||||||||||||||||||||||||||||||||||||||||
Db        301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGC   360

Qy        361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGA   420
              |||||||||||||||||||||||||||||||||||||||||
Db        361 CTAATCAATGCTAAAGCAGCATTAGCTTCACTACAAGA   420

Qy        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATA   480
              |||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA   480

Qy        481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGAAT   540
              ||||||||||||||||||||||||||||||||||||||||
Db        481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATCACGAT   540

Qy        541 AAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCA   600
              |||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATCAAGAGCACTAAAAACC   600
```

```
Qy        181 ATAAATATATCTATCTATGAATAGAAAAGAATCATCTTAAAGACTATTAGTTG   240
Db        181 ATAAATATATCTATCTATGAATAGAAAAGAAATCATCTTAAAGACTATTAGTTG   240

Qy        241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTGAT   300
Db        241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTGAT   300

Qy        301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAGACGCGGAATGCGCGAACAGAT   360
Db        301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGACTATGCCAAGATTGAAGCT   360

Qy        361 CTAATCAATGCTAAAGCTGTAAAGACATTAGCTTCACTACAAGAT   420
Db        361 CTAATCAATGCTAAAGCTAAAGACATTAGCTTCACTACAAGAT   420

Qy        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA   480
Db        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA   480

Qy        481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAACAGCTAATACGGAT   540
Db        481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAACAGCTAATACGGAT   540

Qy        541 AAAAACGACTTTTGATAATGAACAACACCCAAATTTAGTTGATGATGAACCTGAT   600
Db        541 AAAAACGACTTTTGATAATGAACAACAGTTAATAACTAGAGTGTTAGGACTAAAACGT   660
Db        661 CGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACTAGAT   720
Db        721 CCACTAAATGGGGAACGCTTTAGATTCTTTAGATTCTCAATGCTCAATGTCTATCAATTATCTAATAGT   840
Db        841 TTTATTAAAAAGTGATTCAAAATAATGACTGATTTGTAGGGACTTTTACAAAAGAGGACCGTTTGGAATTGGTGATGAACCTTCAAGTAGAATTC   1015
```

```
RESULT 9
S65869
LOCUS       S65869    1015 bp    DNA    linear    BCT 16-MAR-2000
DEFINITION  TM-1=29 kda polypeptide [Mycoplasma gallisepticum, S6, ATCC 15302,
            Genomic, 1015 nt].
ACCESSION   S65869
VERSION     S65869.1    GI:425376
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 1015)
  AUTHORS   Saito,S., Fujisawa,A., Ohkawa,S., Nishimura,N., Abe,T., Kodama,K.,
            Kamogawa,K., Fujisawa,A., Aoyama,S., Iritani,Y. and Hayashi,Y.
  TITLE     Cloning and DNA sequence of a 29 kilodalton polypeptide gene of
```

```
Db   541  AAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181  ACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAACGCACTA  240

Qy   601  ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT  660
          |||||||||||||||||||||||||||||||  |||||||||||||| ||||||||||
Db   595  AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT  654

Db   601  ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Qy   241  AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT  300

Qy   661  CGCAATAATTAGTTGGATCTATACAATAAAGTAGTAGTTAATAACTAGACACTAGAT  720
          ||||||||||||||||||||||||||||||||||||  |||||||||| |||||||||||
Db   655  CAAATTCGCAATAATTAGTTGGATCTATACAATAAAGTAGTAGTTAATAACTAAAAACA  714

Db   661  CGTAATAATTAGTTGGATCTATACAATAATGCTAGTAGTTAATAACTAGACACTAGAT  720
          ||||||||||||||||||||||||||||||||  ||||||||||  ||||||||||
Qy   301  CAAATTCGCAATAATTAGTTGGATCTATACAATAATAACAATAAAGCTAGTAGTTAATAACTAAAAACA  360

Qy   721  CCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT  780
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   715  CTAGATCCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGG  774

Db   721  CCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT  780
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Qy   361  CTAGATCCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGCTAATAAG  420

Qy   781  AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT  840
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   775  AATATTAATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCT  834

Db   781  AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT  840
          |||| ||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Qy   421  AATATCCAGAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCT  480

Qy   841  TTTATTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTTTACAAACGCT  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835  AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTTTACA  894

Db   841  TTTATTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTTTACAAACGCT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481  AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTTTACA  540

Qy   901  AATGTTCAACCTTCAAACTACAGTTTGTTGTTTAGTGCTGATGTAACACCCGTCAAT  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   895  AACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGTTTAGTGCTGATGTAACACCC  954

Db   901  AATGTTCAACCTTCAAACTACAGTTTGTTGTTTAGTGCTGATGTAACACCCGTCAAT  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541  AACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGTTTAGTGCTGATGTAACACCC  600

Qy   961  TATAAAATGCAAGAAGAACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC  1015
          |||||||||||||||||  ||||||| |||||||||||||||||||||||
Db   955  GTCAATTATAAAATATGCAAGAAGAACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT  1014

Db   961  TATAAAATGCAAGAAGAACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC  1015
          ||||||||||||||||||  ||||||||||||||||||||||||||||||
Qy   601  GTCAATTATAAAATATGCAAGAAGAACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT  660

Qy  1015  CTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGTCGGACAAAC  1074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  CTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGTCGGACAAAC  720

Db  1075  ACGAAGTACCAAATTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCT  1134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721  ACGAAGTACCAAATTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCT  780

Qy  1135  TATAAGTTGGTTAAAGCAGTCGATGCTAATAACGTTGGATTACAATACAAAATTAAATAAT  1194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  TATAAGTTGGTTAAAGCAGTCGATGCTAATAACGTTGGATTACAATACAAAATTAAATAAT  840

Qy  1195  GGAAATGTTCAACAAGTTGAGTTGGCCACTTCAACTAGTGCAAATAATACTACAGCTAAT  1254
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  GGAAATGTTCAACAAGTTGAGTTGGCCACTTCAACTAGTGCCACAGCTACAGCTAAT  900

Qy  1255  CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT  1306
          |||||||||||||||||||||||||||||||||||||||||||||
Db   901  CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT  952
```

RESULT 11
E02348
LOCUS   E02348   853 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION   DNA sequence coding for TMG-1.
ACCESSION   E02348
VERSION   E02348.1 GI:2170583
KEYWORDS   JP 190011795-A/7.
SOURCE   Mycoplasma gallisepticum
  ORGANISM   Mycoplasma gallisepticum
               Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 853)
  AUTHORS   Kodama,K., Saito,S., Yanagida,N., Kamogawa,K., Iritani,K. and
               Aoyama,S.
  TITLE   FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE
               THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME
  JOURNAL   Patent: JP 190011795-A 7 24-APR-1990;
               NIPPON ZEON CO LTD, SHIONOGI & CO LTD
COMMENT   OS   Mycoplasma gallisepticum
               PN   JP 190011795-A/7
               PD   24-APR-1990
               PF   02-JUN-1989 JP 1989140283

```
Db   541  AAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC  600
```

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660

Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT 660

Qy   661 CGCAATAATTAGTTGGATCTATACAATAAAGTAGTAGTTAATAACTAGACACTAGAT 720

Db   661 CGTAATAATTAGTTGGATCTATACAATAATGCTAGTAGTTAATAACTAGACACTAGAT 720

Qy   721 CCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Db   721 CCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy   781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840

Db   781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840

Qy   841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTTTACAAACGCT 900

Db   841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTGTTGTTTAGTGCTGATGTAACACCCGTCAAT 960

Db   901 AATGTTCAACCTTCAAACTACAGTTTGTTGTTTAGTGCTGATGTAACACCCGTCAAT 960

Qy   961 TATAAAATGCAAGAAGAACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC 1015

Db   961 TATAAAATGCAAGAAGAACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC 1015

RESULT 10
AX665187
LOCUS   AX665187   1082 bp   DNA   linear   PAT 26-MAR-2003
DEFINITION   Sequence 24 from Patent EP1275716.
ACCESSION   AX665187
VERSION   AX665187.1 GI:29290312
KEYWORDS   .
SOURCE   Mycoplasma gallisepticum
  ORGANISM   Mycoplasma gallisepticum
               Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1
  AUTHORS   Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki,Y.
  TITLE   Modified dna molecule, recombinant containing the same thing, and
               uses thereof
  JOURNAL   Patent: EP 1275716-A 24 15-JAN-2003;
               Zeon Corporation (JP)
FEATURES   Location/Qualifiers
  source   1..1082
               /organism="Mycoplasma gallisepticum"
               /mol_type="unassigned DNA"
               /db_xref="taxon:2096"
               /note="Modified TTM-1 portion (downstream of BglI) of
               pNZ40K-S"

ORIGIN

Query Match   70.6%; Score 921.6; DB 6; Length 1082;
Best Local Similarity 98.0%; Pred. No. 1.4e-124;
Matches 933; Conservative   0; Mismatches   19; Indels   0; Gaps   0;

Qy   355 ACAGATCTAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGACTATGCCAAGAAT 414

Db   1 ACAGATCTAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT 60

Qy   415 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA 474

Db   61 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTGCAGCA 120

Qy   475 ACATTAGAACAACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACAAGCTAAT 534

Db   121 ACATTAGAACAACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACAAGCTAAT 180

Qy   535 ACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 594
```

```
PR    02-JUN-1988 JP 88P  136343
PI    KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,
PI    IRITANI KOICHI, AOYAMA SHIGEMI
PC    C07K13/00,(C12N1/21,C12R1:19),(C12N15/31,C12R1:35),(C12P21/02, PC
      C12P21/02,
PC
G01N33/569,(C12N1/21,C12R1:19),(C12N15/31,C12R1:35),(C12P21/02, PC
      C12R1:19);
CC    strandedness: Double;
CC    topology: Linear;
CC    hypothetical: No;
CC    anti-sense: No;
CC    *source: strain=S6;
FH    Key             Location/Qualifiers
FH
FT    5'UTR           1..39
FT    CDS             40..825
FT                    /product='TMG-1'
FT    3'UTR           826..853.

FEATURES            Location/Qualifiers
     source         1..853
                    /organism="Mycoplasma gallisepticum"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:2096"

ORIGIN

Query Match      63.1%;   Score 824.2;  DB 6;  Length 853;
Best Local Similarity  97.9%;  Pred. No. 2e-110;
Matches  835;  Conservative   0;  Mismatches  18;  Indels   0;  Gaps   0;

Qy  163 TTAGTCATCTTTAAGATATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATC 222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TTAGTCATCTTTAAGATATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATC 60

Qy  223 TTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGT 282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGT 120

Qy  283 ATGTCTATTACTAACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCG 342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ATGTCTATTACTAACTAAAAAGACGCAAACCCAAATAATGGCCAAACCCAATTACAGCAGCG 180

Qy  343 CGAATGGAGTAACAGCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGAC 402
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CGAATGGAGTAACTGATCTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAAGAC 240

Qy  403 TATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TATGCTAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 300

Qy  463 AACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATC 522
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AACCTTAATGCAACACTAGAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATC 360

Qy  523 AACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCA 582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCA 420

Qy  583 TACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCA 642
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCA 480

Qy  643 ACTGCTTATATCAAATTCGCAATAATTAGTGGATCTATACAAATAAAGCTAGTAGTTTA 702
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ACTGCTTATATCAGATTGTAATAATTAGTGGATCTATACAAATAAATGCTAGTAGTTTA 540

Qy  703 ATAACTAAAAACACTAGAACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACT 762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ATAACTAAAAACACTAGAACTAGATCCACTAAATGGGGAATGCTTTTAGATTCTAATGAGATTACT 600

Qy  763 ACAGTTAATCAATGGAATATTAATAATACGTTATCAACTATTAATCAACTAATGCT 822
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ACAGTTAATCAATGGAATATTAATAATACGTTATCAACTATTAATCAACTAATGCT 660
```

```
Qy  823 GATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTA 882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTA 720

Qy  883 GGGACTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTGTTTTAGTGCT 942
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GGGACTTTACAAACGCTAATGTTCAACCTTCAAACTACACAGTTTGTGTTTTAGTGCT 780

Qy  943 GATGTAACACCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCT 1002
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATGTAACACCGTCAATTATAAATATGCAAGAAGGACCGTTTGAAATGGTGATGAACCT 840

Qy 1003 TCAAGTAGAATTC 1015
            |||||||||||||
Db  841 TCAAGTAGAATTC 853
```

```
RESULT 12
E02342
LOCUS       E02342                   708 bp    DNA       linear      PAT 29-SEP-1997
DEFINITION  DNA sequence coding for MG-1.
ACCESSION   E02342
VERSION     E02342.1  GI:2170577
KEYWORDS    JP 1990111795-A/1.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 708)
  AUTHORS   Kodama,K., Saito,S., Yanagida,N., Kamogawa,K., Iritani,K. and
            Aoyama,S.
  TITLE     FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE
            THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME
  JOURNAL   Patent: JP 1990111795-A 1 24-APR-1990;
            NIPPON ZEON CO LTD, SHIONOGI & CO LTD
  COMMENT   OS   Mycoplasma gallisepticum
            PN   JP 1990111795-A/1
            PD   24-APR-1990
            PF   02-JUN-1989 JP 1989140283
            PR   02-JUN-1989 JP 88P  136343
            PI   KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,
            PI   IRITANI KOICHI, AOYAMA SHIGEMI
            PC   C07K13/00,C07K7/06,C07K15/04,C12N1/21,C12N15/31//A61K37/02, PC
                 C12P21/02,
            PC
G01N33/569,(C12N1/21,C12R1:19),(C12N15/31,C12R1:35),(C12P21/02, PC
                 C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   *source: strain=S6;
            FH   Key             Location/Qualifiers
            FH
            FT   CDS             1..708
                                 /product='MG-1'.

FEATURES            Location/Qualifiers
     source         1..708
                    /organism="Mycoplasma gallisepticum"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:2096"

ORIGIN

Query Match      52.1%;   Score 679.8;  DB 6;  Length 708;
Best Local Similarity  97.6%;  Pred. No. 1.9e-89;
Matches  690;  Conservative   0;  Mismatches  17;  Indels   0;  Gaps   0;

Qy  280 TGTATGTCTATTACTAACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCA 339
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TGTATGTCTATTACTAACTAAAAAGACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCA 60

Qy  340 GCGCGAATGGAGTAACAGCTAATCAATGCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAA 399
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCGCGAATGGAGTAACTGATCTAATCAATGCTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAA 120
```

```
Qy    400  GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC  459
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GACTATGCTAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC  180

Qy    460  AATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCC  519
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AATAACCTTAATGCAACACTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCC  240

Qy    520  ATCAACCAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTAGTTGAA  579
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACATCCAAATTTAGTTGAA  300

Qy    580  GCATACAAAGACTAAAAACCACTTTAGAACAAGCTGCTACTAACCTTGAAGGTTTGTCA  639
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  GCATACAAAGACTAAAAACCACTTTAGAACAAGCTGCTACTAACCTTGAAGGTTTAGCT  360

Qy    640  TCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGT  699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCAACTGCTTATAATCGATTCGTAATAATTTAGTGGATCTATACAATAATGCTAGTAGT  420

Qy    700  TTAATAACTAAAAACACTAGATCCACTAAAATGGGGAACGCTTTTAGAGTTCTAATGAGATT  759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  TTAATAACTAAAAACACTAGATCCACTAAAATGGGGAATGCTTTTAGATTCTAATGAGATT  480

Qy    760  ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAAT  819
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAAT  540

Qy    820  GCTGATGCATTATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTT  879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GCTGATGCATTATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTT  600

Qy    880  GTAGGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGT  939
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GTAGGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGT  660

Qy    940  GCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTG  986
           |||||||||||||||||||||||||||||||||||||||
Db    661  GCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGAACGGTTTG  707
```

```
RESULT 13
AE016967
LOCUS       AE016967             301042 bp    DNA     linear   BCT 03-SEP-2003
DEFINITION  Mycoplasma gallisepticum strain R section 1 of 4 of the complete
            genome.
ACCESSION   AE016967 AE015450
VERSION     AE016967.1  GI:31541048
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum R
  ORGANISM  Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 301042)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE     The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
   PUBMED   12949158
REFERENCE   2  (bases 1 to 301042)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzell,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathobiology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eaglevile Road U-89, Storrs, CT
            06269-3089, USA
FEATURES             Location/Qualifiers
     source          1..301042
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
```

```
     rep_origin      1..1889
                     /note="putative oriC"
     protein_bind    complement(24..32)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     pulmonis consensus ttatcMaMa"
                     /bound_moiety="DnaA"
                     complement(36..44)
                     /function="initiation of the chromosome replication"
     protein_bind    /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     pulmonis consensus ttatcMaMa"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     protein_bind    114..122
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     repeat_region   209..217
                     /note="oriC imperfect repeat"
     protein_bind    complement(254..262)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     mycoides consensus YtWtMHaVR"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     protein_bind    298..306
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     mycoides consensus YtWtMHaMR"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     repeat_region   complement(401..450)
                     /note="oriC imperfect repeat"
     repeat_region   complement(462..510)
                     /note="oriC imperfect repeat"
     repeat_region   complement(575..623)
                     /note="oriC imperfect repeat"
     repeat_region   complement(713..761)
                     /note="oriC imperfect repeat"
     repeat_region   complement(763..811)
                     /note="oriC imperfect repeat"
     repeat_region   complement(821..869)
                     /note="oriC imperfect repeat"
     repeat_region   872..920
                     /note="oriC imperfect repeat"
     repeat_region   991..1004
                     /note="oriC imperfect repeat"
     repeat_region   1061..1109
                     /note="oriC imperfect repeat"
     repeat_region   1147..1195
                     /note="oriC imperfect repeat"
     protein_bind    complement(1185..1192)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     capricolum consensus ttWtMHaa"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     repeat_region   complement(1230..1278)
                     /note="oriC imperfect repeat"
     repeat_region   1326..1374
                     /note="oriC imperfect repeat"
     repeat_region   complement(1551..1564)
                     /note="oriC imperfect repeat"
     repeat_region   complement(1609..1657)
                     /note="oriC imperfect repeat"
     repeat_region   complement(1700..1748)
                     /note="oriC imperfect repeat"
     repeat_region   complement(1831..1879)
                     /note="oriC imperfect repeat"
     gene            1890..2661
```

```
CDS
          /gene="parA"
          /locus_tag="MGA_0619"
          /note="synonym: soj"
1890. .2681
          /gene="parA"
          /locus_tag="MGA_0619"
          /note="MGA_0619 ATPase involved in chromosome
          partitioning; pfam00991 ParA family ATPase; COG1192 Soj
          ATPases involved in chromosome partitioning, MGR_001"
          /codon_start=1
          /transl_table=4
          /product="ParA/Soj"
          /protein_id="AAP56351.1"
          /db_xref="GI:31541049"
          /translation="MMIISFINNKGGVLKTTLATNICGVFSKFFPKSRSVIVDLDGQG
          NVSATFGOHPERLKNTLIDIFRGEKDIDDCVLNVFPSIDILPSNHELSFVDMDVARKE
          YKLSVIKNLIEKLEEMYDFVFLDTPPAMSTIVSVSMHEISDMIVIPFEPDQYSILGLLR
          VIDTIETFRERNEDLKVLVVPTKVNARTRLENDVLNIVKTKLSKKNIPLSKNIVSFTT
          KSSASVGVEKLPIVLINQRSKYQBEYIEITKEIINLLKDNKNNEK"

gene
2659. .3147
          /gene="parA/Soj"
          /locus_tag="MGA_0621"

CDS
2659. .3147
          /locus_tag="MGA_0621"
          /note="MGA_0621 PS00318; MGR_002"
          /codon_start=1
          /transl_table=4
          /product="unique hypothetical"
          /protein_id="AAP56352.1"
          /db_xref="GI:31541050"
          /translation="MTRTMKNKKAKKKERRFTDLSADLDEBVEKIDPEYEDFKEIKIE
          KNKDNQVIDKNDPFFYSESFEEARIQLIKDKKVEVKEEBKVQEETVKNKLSEAKKEE
          AKDVYIDSSLEIASQEPLTKGMHFYTNSRIIRKVRECAKNKGLSISRLITMILDKSIK
          EB"

gene
3163. .4548
          /gene="dnaA"
          /locus_tag="MGA_0622"

CDS
3163. .4548
          /gene="dnaA"
          /locus_tag="MGA_0622"
          /note="MGA_0622 ATPase involved in DNA replication
          initiation; COG0593"
          /codon_start=1
          /transl_table=4
          /product="DnaA"
          /protein_id="AAP56353.1"
          /db_xref="GI:31541051"
          /translation="MKTKLKRFLEEISVHFNEANSELLDAFVHSIDFVFBENDNIYIY
          FESPYFNEFKNKLNHLINVENAVVFNDYLSLEWKKIIKENKRVNLLNKKEADTLKEK
          LATLKKQEKYKINPLSKGIKEKNFGNYLVFBFNKEAVYLAKQIANKTTHSNWNPIII
          EGKPGYGKSHLLQAIANERQKLFPEEKICVLSSDDFGSEFLKSVIAPDPTHIESFKSK
          YKDYDLLMIDDVQIISNRPKTNETFFTIFNSLVDQRKTIVITLDCKIEEIQDKLTARM
          ISRFQKGINVRLINQPNKNEIIQIFKQFKENNLEKYMDDHVIEBISDFDEGDIRKIEG
          SVSTLVFMNQMYGSTKTKLIKSPIRKVTMKNLIISKDPKYVFDKIKYHFNVSEDV
          LKSSKRKEIVQARHICMYVLKNVYNKNLSQIGKLLRKDHTTVRHGIDKVEBELENDP
          NLKSFLDFKN"

gene
4744. .5133
          /locus_tag="MGA_1322d"

CDS
4744. .5133
          /locus_tag="MGA_1322d"
          /note="Q COG1132 ABC-type multidrug/protein/lipid
          transport system, ATPase component; similar to MGA_1284;
          MGR_004"
          /codon_start=1
          /transl_table=4
          /product="ABC transporter component"
          /protein_id="AAP56354.1"
          /db_xref="GI:31541052"
          /translation="MNRKLDHCISALLYFGLINFGIIGRDRKNALTQYSLILLGVAAVG
          FTFGYLGGRSIIIASVEFAKQLRVNIPERYQSFSVKNTDKFEKASVLTRMTTDINFIH
          QSIQSGRTAIRGMSVFLFSLVLMFVIS"
```

```
CDS
          /note="synonym: mldB"
5130. .6077
          /locus_tag="MGA_0625"
          /note="MGA_0625, MdlB-like; ABC-type
          multidrug/protein/lipid transport system MdlB [Q] COG1132
          PS00890; similar to MGA_1285, MGA_0626 and MGA_1287;
          MGR_005"
          /codon_start=1
          /transl_table=4
          /product="MdlB-like"
          /protein_id="AAP56355.1"
          /db_xref="GI:31541053"
          /translation="MKLGIASVAIMPVIIGGILLVYRFVIGNYKKLFKQYDQLNNLAK
          BSIAGARVVKSYHQQDNEIOKPNRVAGFIYKNFTKIERTALISPIVLFCIYALAIAI
          AWIGTNNIVDGKLDIGSLASVFAYAFQMLINLLLSVVYVTIITAKPSKDRIIEVLTB
          KIDIKDKKYADTVSDYEVEYKDVSFKYVDTNPHRNLEKINIKIKKGQTIGIIGSTGS
          GKTSIVNLLTRLYECNEBGQVLLNNIQLNNYSIKALRDAIAIVPQKSILYSGTIKDNIL
          MGGNYSDEEVBKAITQAQAAEFINKLPNKLDSIVBQKWN"

gene
6369. .8294
          /locus_tag="MGA_0626"

CDS
6369. .8294
          /locus_tag="MGA_0626"
          /note="MGA_0626, ABC-type multidrug/protein/lipid
          transport system, similar to MdlB [Q] COG1132 PS00890;
```

```
Query Match          19.2%;   Score 250.6;   DB 1;   Length 301042;
Best Local Similarity 55.3%;   Pred. No. 5.3e-28;
Matches  587;   Conservative  0;  Mismatches  444;   Indels  30;  Gaps  4;

Qy    251 CATCTTCTTAGCATTGGCAAACCAATTAGAAGCAGCGCGAATGGAGTTAACAGATCAATCAATCAATG 310
                |||  ||||||| || ||  ||||| |||||| ||||||||||| ||||||||
Db 258430 CAAACCCTCCTAGTGGTGGTATGATGTCGCGGAGATACTAATCCAGGAAATGACGGAGAA 258489

Qy    311 CAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATCAATCAATG 370
                ||||  |||||| | ||||||||||||||| |||||||
Db 258490 TGGAGAATTCTGCTCAACAAGTAGCAGCTGCTAAAATCAAGAATTAAGCGATTTATTAGCAA 258549

Qy    371 CTAAAGCCATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAACGTAGTTTATCAT 430
                |  |||||| |||||||| ||| |||| ||||||
Db 258550 CACAAAGCAGTAACCTTGCAAAATATGCAGATTATCAAATATTCAGGAAATACTCTAACAG 258609

Qy    431 CTGCTTATAGTGAAGCTGAAACAGTTAACATTAACCTTAATGCAACATTAGAACAACTAA 490
                | |||||||||||| ||||||||| |||| || ||||
Db 258610 CTGCATACACAACTACTACTACATGCAGATTATCAGATAATGACAAGCGTCACTTTAGGCAGTTA 258669

Qy    491 AAATGGCTAAAACTAATTAGAATCAGCCATCAACAAGCTAATACGGATAAAACGACTT 550
                |||| || |||| ||||||||||||  || |||| |||||||
Db 258670 AAAGTGCTACATCAACTTACAGCAGCTATTGATCTGCTGCAAGTTCAAAAACTAGTT 258729

Qy    551 TTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAACCACTTTAGAAC 610
                || || |||||| | ||||| || ||| | ||||||
Db 258730 TTGATGAAAAAAATCCTGAATTAATCAAAGCATATTATGCTTTAAAAGAAACATTAAAAA 258789

Qy    611 AACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATAATCAAATTCGCAATAATT 670
                || ||||||| |||||||| ||||||||  || ||||||||
Db 258790 ATGAAGAACAGTTTAAGTGGTTAACTGATTCTAATTTTGCAACAATCAAAAACTAACC 258849

Qy    671 TAGTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAAACACTAGATCCACTAAATG 730
                || |||||| || |||||||| |||||||||||||||
Db 258850 TTACTGCTTTATATCAATCTGGTAAAGATTTTGTTAAGCAACACATTAGATCATCCTGT---TA 258906

Qy    731 GGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATTAATAATACGT 790
                ||  |||||||||  |||||||||||||||| |||||||
Db 258907 GTGGAAACGCTCCCCAAGTGCTGATATTGATATCGT 258966

Qy    791 TATCAACTATTAATGAACAAAGACTAATGCTGATGCATTACTAATAGTTTTATTAAAA 850
                ||||||  ||||||||| || |||| |||||||
Db 258967 TTTCAAAACTTGAAACGTGAAAAAACAAATGCTAATACTTTAGCTACAAGTTTTGTTAAAG 259026

Qy    851 AAGTGAATCAAAATAATGAACAAAGTTTGTAGGGACTTTACAAACGCTAATGTTCAAC 910
                ||||||| || |||| |||| ||||||  ||||||
Db 259027 AGTTTTAGTTAAAATAAATTAACAGGAATTGATAC---AACTAATAATCGAGAACAAAC 259083
```

```
Qy     911  CTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAATTAT-------  963
            ||||||  ||||  |||| ||| ||||||| | || |||| |||||||||||
Db  259084  CTGGAAATTACAGCTTTGTTGGTTATAGTGTTAATGCAACAAATAATAATGAGATTCCTA  259143

Qy     964  ----AAAATGCAAGAAGGACCGTTTGGAAATGGTGAAGTCAAGTAGAGAATTCTTG  1018
                ||||||  ||| || |||  ||||||||| ||||  ||||| ||||| ||||
Db  259144  ACTGGAAATTTCGCACAAAGAAAGGTTTGAACTAGTAGACAGTAGAACTAGTCTAATTT  259203

Qy    1019  CAAACACGAATAGTA-----------TCACAGATGTTCTTGGATTTATAGTTTAGTCTG  1066
            |||  ||| ||  |             |  || ||| || ||| |||| ||||| |||
Db  259204  CTAGCCACTTCCTGATAATTCTCTCCACCTTAACAGAAGTATCATGATGATTTATAGTTTATCAG  259263

Qy    1067  GAACAAACACAGAAGTACCAATTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTAT  1126
            ||| |||| |||||| || ||||||| || |||||| ||| |||||||||||| |||||
Db  259264  GAGCAGGTACTAAATACAGCTTAACATTTAATTAATTACGGTCCGTCCAACTGGTTATTTAT  259323

Qy    1127  ATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACATACACAAAT  1186
            ||||||| ||||||| ||||| || ||| |||| |||||| ||| |||||  |||||||
Db  259324  ATTTCCCTTATAAGTTAGTTAAACAAGGTGATGAAAATAATGTAGCTTTACAATATACCT  259383

Qy    1187  TAAATAATGAAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAAATAATACTA  1246
            |||||||||||| |||||||||||| || ||| |||| ||||||||||  |||
Db  259384  TAAAATGGTGGTAGTGCACAAGAAGTTAAGTTGCTCCAACTGTAAAAACTAGTGTTAGTG  259443

Qy    1247  CAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCT  1287
            |||| ||| ||||| ||||  |||||||||  ||||||||
Db  259444  CAGATTCTTCTGGTGATTCAAATAATCAGACAGAAGTGCT  259484
```

```
RESULT 14
AF210770
LOCUS       AF210770                4568 bp    DNA     linear   BCT 02-FEB-2000
DEFINITION  Mycoplasma gallisepticum pMGA-like protein 9.1 gene, partial cds;
            pMGA-like protein 9.2 gene, complete cds; and pMGA-like protein 9.3
            gene, partial cds.
ACCESSION   AF210770
VERSION     AF210770.1  GI:6851355
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 4568)
  AUTHORS   Pharr,G.T., Branton,S.L., Hanson,L.A., Minion,F.C., Lott,B.D.,
            May,J.D. and Hughlett,M.B.
  TITLE     A novel pMGA-like gene from the F-strain (vaccine strain) of
            Mycoplasma gallisepticum
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4568)
  AUTHORS   Pharr,G.T., Branton,S.L., Hanson,L.A., Minion,F.C., Lott,B.D.,
            May,J.D. and Hughlett,M.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1999) College of Veterinary Medicine, Mississippi
            State University, Box 9825, MS 39762, USA
FEATURES             Location/Qualifiers
     source          1..4568
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="F"
                     /db_xref="taxon:2096"
     CDS             <1..637
                     /note="adhesin"
                     /codon_start=2
                     /transl_table=4
                     /product="pMGA-like protein 9.1"
                     /protein_id="AAF29523.1"
                     /db_xref="GI:6851356"
                     /translation="VYNDIFSNILDKQDEPNSVTVNLLKGYSLAADRSTYFYQFSSSN
                     RMNESNPTYLVGFIGGHGNRNNLNSSNTTNNEVASPSVQTSNRTLITIYNAPKDGQYY
                     IKGSVLTSNNRNLKFTTTATANNSITFTVKGKNNWSTLGTFNTANNNDIETSGSSSSG
                     QPNESKTIKLNKGLNKVVITSVMNDNKNPGAPYIGNLIFTLMSPTMEAKK"
     repeat_region   787..831
                     /note="microsatellite"
                     /rpt_type=tandem
```

```
CDS             967..2982
                /note="adhesin"
                /codon_start=1
                /transl_table=4
                /product="pMGA-like protein 9.2"
                /protein_id="AAF29524.1"
                /db_xref="GI:6851357"
                /translation="MRKNIIKFVSLLGIGSFVMLAAASCTSAATPTPNPEPKPDPMP
                NPPSGDMNGGDTNPGNDGGMENSAQQLAAAKKELSDLLATQSSNLAKYADYTNIQNTL
                TAAYTTAKSTSDNTSVTLEQVKSATSTLQAAIDTRAASKTSFDEKNPELIKAYYALKE
                TLKNEETVLSGLTDSNFATIKTNLITALYQSGGDFVKATLDPVSGNAPQIADITKADKD
                IADAVSKLETWKTNANTLATSFVKVLVKNKLTGIDTTNREQPGNYSFVGYSVNATN
                NNEIPNWNFAQRKVWTSGDNGRTSLISSTSDNSSTLTEVSWIYSLSGAGTKYSLTFNYY
                GPSTGYLYFPYKLVKSGDENNVALQYTLNSGSAQEYNFAPTVKTSVSADSSGDSNNQT
                ESAAETMPVTSDLNPATVSDINIAKLTLSNLKFGSNTIEFSVPTEPSNKVAPMIGNM
                YLTSNIANEAKVYNSIFGNVDNSSEASTYVTVDILKGYSLATNWSTYVTRFMNLTNSM
                PENATTYLVGFIGGQLARTTVGSIPNRNNFPIMNNENRTFLYVNAPKAGDYHISGSY
                LTRNTRGLKLIVTDTTDKNNSITITTSGKNWNTLGHPDTSKANNSNGNDGSVENNKA
                SLTLKEGLNKIVLIAGGTQDGKNAPYIGNLIFTLNNSSTNASQDSSST"
repeat_region   3147..3182
                /note="microsatellite"
                /rpt_type=tandem
                /rpt_unit="gaa"
CDS             3313..>4568
                /note="adhesin"
                /codon_start=1
                /transl_table=4
                /product="pMGA-like protein 9.3"
                /protein_id="AAF29525.1"
                /db_xref="GI:6851358"
                /translation="MRKNIIKFVSLLGIGSFVMVAAASCTSAATPTPNEPNPNEP
                KPDPMNPPSGGMNGGDINPGGGQNMMDSAAQELTAARTALTSLLASKNANIEMYSDY
                AKIQNTLIAAYTTAEQTSQNSSATLEQVKNATSALQTAINTANSNKQKFPQDHSNLLM
                SYKNLMATLAKKETTVMTLKDPKYSAILDQINGVSCKGBELVQHTLDPVSGIVPAANT
                ITEBITKIEBVISEKTLQDQKNNADQFDNYQSPTLDKTKLENVEDAKKMGQPANYSFV
                GYSVDVTGSGQETTIPNWNFAQRAIFTSGNQPTKVTATTTGEDQSTAKPLSDVSWIY
                SLAGTGAKYTLSFTYYGPSTGWLYFPYKLVXANDDVGLQYKLNSNETILTPIIFGEGTT
                TNGPAATVENINVAKVRLTGLAFGKNT"
```

```
ORIGIN
        Query Match            19.1%;  Score 249;  DB 1;  Length 4568;
        Best Local Similarity  55.2%;  Pred. No. 3.5e-27;
        Matches   586;  Conservative   0;  Mismatches  445;  Indels  30;  Gaps  4;

Qy     251  CATCCTTCTTAGCATTGGGATTTCTAGCTGTGTATGTCTCATTACTCATTAAAAAGATGCAAACC  310
            |||||  ||| |||||| |||| |||||  | || || | |||| ||| |||  |||||||||
Db    1097  CAAACCCCTCCTAGTGGTGATATGAAATGGCGGAGATACTAATCCAGGAAATGACGGAGGAA  1156

Qy     311  CAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTCAACATCAGATCTAATCAATG  370
            |||| ||||| |||| ||||| ||||||| ||| ||||||  ||||||| |||| |||||||
Db    1157  TGGGAGAATTCGTCTGCCAACAATTAGCAGCTGCTAAAAAAGAATTAAGCGGATTTATTAGCAA  1216

Qy     371  CTAAAGCGATGACATTAGCTTCCACTACAGAGACTATGCCAAGAGAGATTGAAGCTAGTTTATCAT  430
            | ||||| ||| |||||| |||| ||| ||| ||| | ||||||| |||||||  |||||||
Db    1217  CACAAAGCAGTAACCTTGCAAAATATGCAGATTATACAAATATTCAGAATACTCTAACAG  1276

Qy     431  CTGCTTATAGTGAAGCTGAAACAGTTAACATCTACACATCAGATATAAAAGTTGTGT  490
            |||||| || ||||||| || ||||| |||| || ||||||  ||||  ||||
Db    1277  CTGCATACACAACTGCTAAATCACACTCATCAGATATACATCAGATATTATGATGTGTCACTTTAGGA  1336

Qy     491  AAATGGCTAAAACTAATTTAGAATCAGCCATCAGCCAAGCACTAAAAACCACTTTAGAAC  550
            ||||| ||||| |||| ||||||| ||| || |||||| || ||||||||||| ||||
Db    1337  AAAGTGCTACATCAAGTTTACAGCAGCTATTGACAGCAGCCAAGTTCAAGAAAAACTAGTT  1396

Qy     551  TTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAAC  610
            |||||||||||| | ||||| ||| |||| ||  || ||||||| || |||| ||||||
Db    1397  TTGATGAAAAAAATCCTGAATTAATCAAAGCATTATATGCTTTAAAAGAAACATTAAAAA  1456

Qy     611  AACGTGCTACTAACCTTGGAAGGTTTGTCATCAACTGCTTATAATCAAATCAAAACTACC  670
            ||   ||| || |||  |||||| | ||  || ||||||| ||| ||| ||||||| |||
Db    1457  ATGAAGAAACAGTTTAAGTTGTTTAACTGATTCTAATTTGCAACATCAAAACTACC  1516
```

```
Qy   671  TAGTGGATCTATACAATAAAGCTAGTAGTTTATAATAACTAAAACACTAGATCCACTAAATG  730
             |||| |||| ||
Db  1517  TTACTGCTTTATATCAATCTGGTAAAGATTTTGTTAAAGCAACATTAGATCCTCGT---TA  1573

Qy   731  GGGGAACGCTTTTAGATTCTAAATGAGAATACTACAGTGAATATTAATAATACGT  790
             ||| |||
Db  1574  GTGGAAACGCTCCCCAAATGCTGATATTACAAAAGCAGATAAAGATATTGCTGATGCTG  1633

Qy   791  TATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGTTTTATTAAAA  850
             ||| ||| |||
Db  1634  TTTCAAAACTTGAAACCTGAAAAAACAAATGCTAATACTTAGCTACAAGTTTGTTAAAG  1693

Qy   851  AAGTGATTCAAAATAATGAACAACAAAGTTTGTGAGGACTTTACAAACGCTAATGTTCAAC  910
             | ||| ||| |||
Db  1694  AGTTTTAGTTAAAATAAATAACAGGAATTGATAC---AACTAATAATCGAGAACAAC  1750

Qy   911  CTTCAAACTACAGTTTGTTGTGCTTTTAGTGCTGATGTAACACCCGTCAATTAT-------  963
             |||| ||| |||
Db  1751  CTGGAAATTACACGCTTGTTGGTTATAGTGTTAATGCAACAAATAATAATGAGATTCCTA  1810

Qy   964  -----AAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTG  1018
             ||| ||| |||
Db  1811  ACTGGAATTTCGCACAAAGAAAGTTGAACTAGTGATAACGGTAGAACTAGTCTAATTT  1870

Qy  1019  CAAACACCGAATAGTA---------TCACAGATGTTTCTTGGATTTATAGTTTAGCTG  1066
             ||| |||
Db  1871  CTAGCACTTCTGATAATTCTTCTACCTTAACAGAAGTATCATGAATTTATAGTTTATCAG  1930

Qy  1067  GAACAAACCGAGTACCAATTTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTAT  1126
             ||| ||| |||
Db  1931  GAGCAGGTACTAAATACAGCTTAACATTTAATTATTACGGTCCTTCACTCGGTTATTTAT  1990

Qy  1127  ATTTCCCTTATAAGTTGGTTAAAGCAGCTGATCGTAATAACGTTGGATTACAATACAAAT  1186
             ||| ||| |||
Db  1991  ATTTCCCATATAGTTAGTTAAAGAAGGTGATGAAAATAATAATGTAGCTTTACAATATCCT  2050

Qy  1187  TAAATAATCGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTA  1246
             ||| ||| |||
Db  2051  TAAATAGTGGTAGTGGCACAAGAAGTTAATTTTGCTCCAACTGTAAAAAACTAGTGTTAGTG  2110

Qy  1247  CAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCT  1287
             ||| |||
Db  2111  CAGATTCTTCTGGTGATTCAAATAATCAGACAGAAAGTGCT  2151
```

```
LOCUS       301903 bp   DNA   linear   BCT   03-SEP-2003

source          1..301903
                /organism="Mycoplasma gallisepticum R"
                /mol_type="genomic DNA"
                /strain="R"
                /db_xref="taxon:233150"
gene            complement(77..754)
                /locus_tag="MGA_0090"
CDS             complement(77..754)
                /locus_tag="MGA_0090"
                /note="MGA_0090; [R] COG1881 Phospholipid-binding protein;
                MGA_456"
                /codon_start=1
                /transl_table=4
                /product="predicted phospholipid-binding protein"
                /protein_id="AAP56784.1"
                /db_xref="GI:31541484"
                /translation="MRIMKYSIGNNSKIWSKMVGADGYLALHGGGHKNEQTPPSF
                SLDLSWEKIPNAKSYAVIVEDFDAAAVIGFSFIHWIAVNIKTNHIAENQSYLDHNKWL
                ESKQSAYGENLLWQGHNSSVNKTLVANNKTNKDLGGILPEGFTSTSLENSLMYFGCYP
                PNADHTYTVVYGLDVDASLTYHKNHRTQKYHLNKPYYVGDFMQAIHKHVGKHTLH
                FKYKKVG"
gene            complement(767..1396)
                /locus_tag="MGA_0091"
CDS             complement(767..1396)
                /locus_tag="MGA_0091"
                /note="MGA_0091; [R] COG1881 Phospholipid-binding protein;
                MGA_457"
                /codon_start=1
                /transl_table=4
                /product="predicted phospholipid-binding protein"
                /protein_id="AAP56785.1"
                /db_xref="GI:31541485"
                /translation="MKKYISYIFNLNYYFDNYLRWIINMDKKKHAIWSNCFTETGDG
                VYLKSETVKLNNRYESLDLTWDKIEGAKSYAVMVDYEASRVIGQSFIHWVVANVKE
                NKLAYAANINNKSIVQGLNSTAQCQTESSKGVIIECVFGAFKNTAEAASDYLPPLPPD
                DTHLYTLRVYGLDVDHVDLPKHFSLTDLNQKIIRHTVGEHBLHFWYKPN"
gene            complement(1437..2306)
                /gene="araI"
                /locus_tag="MGA_0093"
CDS             complement(1437..2306)
                /gene="araI"
                /locus_tag="MGA_0093"
                /note="MGA_0093; ARA1 [R] COG0656 Aldo/keto reductases
                related to diketogulonate reductase; MGR_458"
                /codon_start=1
                /transl_table=4
                /product="ARA1"
                /protein_id="AAP56786.1"
                /db_xref="GI:31541486"
                /translation="MYLTTKKIKVMKYNLLANNYKMPSIGFGTYKLEDGKQTVEAVKT
                ALEVGYRHLDCABIYENQLSVSQGIKLSGVNRKEIFITSKIWNDDKGYESTKKAFHKI
                LKDLDLBYLDLLLIHWPIGKGFKDNWQEVNAQTWKAMEEFYLEGKIKAIGLSNFIVHH
                IEALKKSAKILPMVNQLEFHPGYLQPEIVEYCQKNNIVQAWSPFSQSKVLENKTLVQ
                LAKKYEVTVAQLIANWIIQKDIVPLPKSKTPBRIRSNFDVFNFGINQADLKVIDQLND
                CGGLGLHPDQVDF"
gene            2541..2930
                /locus_tag="MGA_0094"
CDS             2541..2930
                /locus_tag="MGA_0094"
                /note="MGA_0094; MGR_459"
                /codon_start=1
                /transl_table=4
                /product="conserved hypothetical"
                /protein_id="AAP56787.1"
                /db_xref="GI:31541487"
                /translation="MINLLIPTMVFNKSKDLLISKPRIYINLKSNMISKLVRVDLNSV
                ILQIBYNWATBRKHKRNLIKKIDRLIDSWNNDLKKVISKPKRLHRWRDBIDYFTSR
                IINLQKNELNLPLNNRRPDQLLARKKK"
gene            complement(2997..3440)
                /locus_tag="MGA_0096"
CDS             complement(2997..3440)
                /locus_tag="MGA_0096"
                /note="MGA_0096; MGR_460"
```

```
                    LKEIKTEFSFIKQITPLEFSKQYVNYQEYDFLDLWTIGNTIDYKINQRYSLKDSNNLF
                    EASDFNFLNVDRNILKFFALANLVAKQTMWPACDVNHYRNNKTGPDNQQPDYQSLFN
                    IDFSVDRNKQVRSHFISSNHAMTLSGVDFPDEAKSLLKQKELVKKFKPDQYQFVL
                    DLSQTPVFKKWKIENSWGEKVGNKGFYYMNDQWFNDYLIDIVISTKAADDFFKNPKPI
                    DPDLIGLVKQLKTKEILKAGLDTKPILIDGYDPLGANNKGVK"
gene    complement(6968..7402)
                    /gene="smpB"
                    /locus_tag="MGA_0102"
CDS     complement(6968..7402)
                    /gene="smpB"
                    /locus_tag="MGA_0102; SmpB [O] COG0691 tmRNA-binding protein;
                    MGR_465"
                    /codon_start=1
                    /transl_table=4
                    /product="SmpB"
                    /protein_id="AAP56793.1"
                    /db_xref="GI:31541493"
                    /translation="MMRLLVNNKAKYNYELLDKYEAGISLSGNEVKSLALHHGKLDD
                    SYVIIRKNEAYLLNLLIPKYKFDTSKVLNETRTRKLLLHKSEILKIDLIKKQHSLVII
                    PYQIYFVNNKIKVSIYLARPKKRYDKRQTIKEREINKKIRKY"
gene    complement(7399..8148)
                    /locus_tag="MGA_0103"
CDS     complement(7399..8148)
                    /locus_tag="MGA_0103"
                    /note="MGA_0103; MGR_466"
                    /codon_start=1
                    /transl_table=4
```

```
Query Match     17.1%;     Score 222.8;     DB 1;     Length 301903;
Best Local Similarity   56.2%;     Pred. No. 5.7e-24;
Matches   516;   Conservative   0;   Mismatches 372;   Indels 30;   Gaps 4;

Qy      323 AAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATCATGCTAAAGCGATGA   382
                 ||   |||| || | |||| ||| |||  || || ||| | | ||||  |||   || ||
Db   220537 AACAACAATTAGCAACTGCTAGAAAAACGCTAACCGATCTAATTGGTACACGAAAATACCA   220478

Qy      383 CATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCTACTAGTTTATCATCTGCTTATAGTG   442
                 | |||  ||||| | |||||| || ||||| || |||  |||| |||  |||| ||||||
Db   220477 ACGTTGCTTTATATGCTGATTATGCCAAATTCAAGCACTTTAAGCACTGCTTATATGA   220418

Qy      443 AAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGACAACTAAAAATGGCTAAAA   502
                 ||||| |||| | || |||| |||  |||||||| |||||  |  ||||  |||| |
Db   220417 CAGCTAAAACTGCATCAGCAAATACAAATGCAACTTTAGAAAATCTAAGATCTCATCAA   220358

Qy      503 CTAATTAGAATCAGCCATCACCAAGCTAATACGGATAAAACGACTTTGATAATGAAC   562
                 |||  |||||| || | ||| | |||||| || |||||| | || ||| |||||||
Db   220357 CTACACTACAAACAGCAATTGATAAGCTGTTAGTGATAAAAATGATTTAATAACCAAA   220298

Qy      563 ACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTA   622
                 |   | ||| ||||||||| || ||||| |||||  | | |||||| | ||| | |||
Db   220297 ATGCTGATTAGTTAGTGCATATACATCATTAAAAGATGCTGTTAAATCAGAAACTATGA   220238

Qy      623 ACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATTCAACATAATTTAGTGGATCTAT   682
                 |  | |||||||| ||| ||| |  |||| | ||| |||| | | ||||| | |||| |
Db   220237 ATTTAGAAGGTTTATCTGATGCAAATTATACAGCCAAATAACCTTACTACTTATT   220178

Qy      683 ACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTT   742
                 | || |||||| | | |||  || |  |||| | || | |||| ||||| ||| |||
Db   220177 ATGGTAAAGCTAAGAATATTTAACAGCTACGCTATATCCTGAATCAGGGGGGAGATTCCTA   220118

Qy      743 TAGATTCTAATAGAATTACTACAGTTAATCGAATTTTATTACAAATGCAAATAGTAATATTG   802
                 ||| |   |||| | | | | ||| |||  |||||||||||||||| ||||||||||||||
Db   220117 AAG--TTGATGCTGTAAAAGTAGTAAACAATTTTATTACAAATGTCAAATAGTAATATTG   220060

Qy      803 ATGAACAAAGACTAATGCTGATGCAATATCGATAGTTTATTAAAAAAGTGAATTCAAA   862
                 | ||||||| | |||||| || |||| |||| |||||| | |||| ||| ||| | |
Db   220060 AACAATGAAAAAATAATGCAAACATGTGATCTGATAGTTGTTTAAAGAAAAACCTTAGATA   220001

Qy      863 ATAATGAACAAAGTTTTGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACA   922
                 |  | |||||| ||| || ||||  | ||||| || |||| || |||  |  ||| ||
Db   219953 AAGAACAATTAACAAGTAATGCGGAATTACAA--------CAACCTGCGAAATTACA   220000
```

```
                    /codon_start=1
                    /transl_table=4
                    /product="unique hypothetical"
                    /protein_id="AAP56788.1"
                    /db_xref="GI:31541488"
                    /translation="MKKPKFKKVVSIKMLNKLAEKLKMTKKGELKERLNVSDDVSESQLL
                    KMAEFMEKLQSMYGDAIFNHDYGAESDDEKEDKPISAPKDKKQNLS"
gene    complement(3397..3609)
                    /locus_tag="MGA_0097"
CDS     complement(3397..3609)
                    /locus_tag="MGA_0097"
                    /note="MGA_0097; MGR_461"
                    /codon_start=1
                    /transl_table=4
                    /product="unique hypothetical"
                    /protein_id="AAP56789.1"
                    /db_xref="GI:31541489"
                    /translation="MSSFLGFVMIGFFIHIKERNNSQDLAYKTNLFLKGHYFFKITK
                    YQIKQLIIAKALIKEIOESSINQDVK"
gene    complement(3483..3872)
                    /locus_tag="MGA_0099"
CDS     complement(3483..3872)
                    /locus_tag="MGA_0099"
                    /note="MGA_0099 unique hypothetical; MGR_462"
                    /codon_start=1
                    /transl_table=4
                    /product="nique hypothetical"
                    /protein_id="AAP56790.1"
                    /db_xref="GI:31541490"
                    /translation="MIRTMRSNWNPSQIIDEINNSIKNKKQPMILILVGFIIGMIVFG
                    IGIYLSLSNPDSNNINRSLKTNDMNTRVENSISINSPKVVFDRFYIHFWAIWWLGPSY
                    TSKNEILKIWLIKQIYSWKNIIIFLK"
gene    complement(3830..5323)
                    /gene="pepC"
                    /locus_tag="MGA_0100"
CDS     complement(3830..5323)
                    /gene="pepC"
                    /locus_tag="MGA_0100"
                    /note="MGA_0100; PepC [E] COG3579 Aminopeptidase C;
                    MGR_463"
                    /codon_start=1
                    /transl_table=4
                    /product="PepC"
                    /protein_id="AAP56791.1"
                    /db_xref="GI:31541491"
                    /translation="MMAKLELKDSLKQLNNLSRSKFNKQLTNILNTVGVKSVSYNGYN
                    FSKNSLSFNLDLSAQPITNQFQTGRCMIFAGLNLLRYHLAKELNIDDLELSQSYLAFW
                    DKPEKANYFLETVTELRDKKVVNDRTLSFILKNGVADGGQWTMLTNLIKKYGLVPKSIM
                    PDTANSASTNQLNYLINLKLNQAASKILDNKDKPVAELNKIKTKALDEVFYLLSSIYG
                    ELPTKPDFSYTKVIKDSADRSBSHFNVLPKKVVETNFTPLIFYQKYFKQLVEKQSFV
                    SVINAPANNKFNQTFSVKYLNNVIESDEILHYNLEQLFSVLTIKQLVNQQPVWFGC
                    EVKNIYLNEAKTFWDDSFDYQTLFNIDLSLSKNDQLDYWLSSINHAMLITGVDLDLE
                    KFNQIDQEFNQLVKANKTKQAYQYLVEHMDQLTIHKWKIENSWGQDLGHAGYFVISDS
                    YFKRYTYQVAINRQLFNQLVEQLKLTKEFDKPVRLLSPWDPIGTLAK"
gene    complement(5320..6972)
                    /gene="pepC"
                    /locus_tag="MGA_0101"
CDS     complement(5320..6972)
                    /gene="pepC"
                    /locus_tag="MGA_0101"
                    /note="MGA_0101; PepC [E] COG3579 Aminopeptidase C;
                    MGR_464"
                    /codon_start=1
                    /transl_table=4
                    /product="PepC"
                    /protein_id="AAP56792.1"
                    /db_xref="GI:31541492"
                    /translation="MKSMKMKNKNKIDFQKYLKKQLSYQKKREYLLTRNALVANDINLM
                    AANVFNPDKNLTEFLVDAHKPTLTITNQDMSGRCWIFAGLNPLRRQTAEKLKVSNFVF
                    SQTYMDFWDKYERANVFLNEIKADVELDDRDLKAELQSAGQDGGWYGFFENLVNKY
                    GLVPQEVMPDSFSGHNTFILNELLQVVLIKATKEIRAHKKASQKQKEVVDATLKKVLE
                    MLVLAYGPVPSKPDWQYVADAKDEENKELANKEEKSTKTKTEVKTPAQQEEEKAKQE"
```

```
Qy       923 GTTTTGTGCTTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCG 982
             ||||||||| ||||| |||  ||| ||||           |  ||||||| ||| ||||
Db    219952 GTTTTGTAGCATATAATCAAGATATAACTAATCCTACTTATAATTTTGCTAAAAGAGTAG 219893

Qy       983 TTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGC--------AAACACGAATAGTA 1033
             |||||| ||| ||    |||||| |||||||||| | ||        |
Db    219892 TTTGAAAACCACAAGAAGGTCGGTCAAGTACCTATGTGCCATTAGAAAATCAAGGTGATT 219833

Qy      1034 TCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAACGAAGTACCAATTTAGTT 1093
             ||  ||||| |||| |||||||||||||| |||||| || ||| ||  ||||||||
Db    219832 TAACAGACGTATCATGAATTTATAGTTTAGCTGGAAATGAAACTAAATATAGTTTTACTT 219773

Qy      1094 TTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAG 1153
             ||  ||| || || || | ||| ||   |||| ||| |  || |   ||||  |||
Db    219772 TTGCTAATTATGGCACACTACAGGATACTTATATTTTCCATACACGTCAGTTAAATCAA 219713

Qy      1154 CTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATAATGGAAATGTTCAACAAGTTG 1213
             ||||| |||||||| || |||| ||||| |||| |||||| |||| ||  ||||| |||
Db    219712 GCG------ATAATGTTGCATTACAAATATAAGTTAAATGGTTGCTAGTCCGGTATCTATCG 219659

Qy      1214 AGTTTGCCACTTCAACTA 1231
             ||||| |||||||||
Db    219658 ATTTTAATAACTCAGCAA 219641
```

Search completed: May 5, 2004, 17:06:39
Job time : 4902.91 secs

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:  May  5, 2004, 09:54:21 ; Search time 457.318 Seconds
                                   (without alignments)
                         12131.925 Million cell updates/sec

Title:          US-09-901-572A-1
Perfect score:  1306
Sequence:       1 aaaaacatccagattgttaat.........taaaacgtttatcaggtt 1306

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:    6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

                 SUMMARIES

Database :   N_Geneseq.29Jan04:*
         1:  geneseqm1980s:*
         2:  geneseqm1990s:*
         3:  geneseqm2000s:*
         4:  geneseqm2001as:*
         5:  geneseqm2001bs:*
         6:  geneseqm2002s:*
         7:  geneseqm2003as:*
         8:  geneseqm2003bs:*
         9:  geneseqm2003cs:*
        10:  geneseqm2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1282	98.2	2346	2	AAQ77857	Aaq77857 Mycoplasm
2	1264	96.8	1387	2	AAQ77854	Aaq77854 Mycoplasm
3	1262.4	96.7	1387	2	AAQ53419	Aaq53419 Mycoplasm
4	1260.8	96.5	1387	2	AAQ94711	Aaq94711 Mycoplasm
5	1086.8	83.2	1935	2	AAQ77856	Aaq77856 Mycoplasm
6	1028	78.7	1152	7	ACF03364	Acf03364 Mycoplasm
7	1024.2	78.4	1371	2	AAT96595	Aat96595 Hybrid Ma
8	1022.0	78.3	3261	2	AAT96596	Aat96596 Hybrid Ma
9	989.4	75.8	1015	2	AAT04076	Aat04076 M.gallise
10	921.6	70.6	1082	7	ACF03382	Acf03382 M. gallis
11	824.2	63.1	853	1	AAQ04687	Aaq04687 DNA encod
12	801.8	61.4	853	1	AAN92574	Aan92574 DNA seque
13	679.8	52.1	708	1	AAN92568	Aan92568 DNA seque
14	678.2	51.9	708	1	AAQ04686	Aaq04686 DNA encod
15	216.6	16.6	2196	2	AAT04075	Aat04075 M.gallise
16	216.6	16.6	2369	2	AAQ77853	Aaq77853 Mycoplasm
17	205.6	15.7	2452	2	AAQ68670	Aaq68670 PMGA1.3 M
18	205.6	15.7	2453	2	AAT51536	Aat51536 Mycobacte
19	183.8	14.1	2416	2	AAQ68671	Aaq68671 PMGA1.4 M
20	183.2	14.0	2417	2	AAT51537	Aat51537 Mycobacte
21	174.8	13.4	2809	2	AAT51531	Aat51531 Mycobacte
22	173.2	13.3	2282	2	AAT51535	Aat51535 Mycobacte
23	173	13.2	2881	2	AAQ68668	Aaq68668 PMGA 1.2
24	160.6	12.3	2004	2	AAT51538	Aat51538 Mycobacte
25	150.8	11.5	2337	2	AAQ68669	Aaq68669 PMGA1.2 M
26	149.2	11.4	2577	2	AAQ68672	Aaq68672 Partial P
27	136.4	11.4	702	1	AAQ05650	Aaq05650 DNA encod
28	136.4	10.4	702	1	AAN92570	Aan92570 DNA seque
29	78.2	6.0	594	1	AAQ05651	Aaq05651 DNA encod
30	78.2	6.0	594	1	AAN92571	Aan92571 DNA seque
31	73.4	5.6	4985	6	ABQ75107	Abq75107 Anopheles
32	73.4	5.6	4985	9	ACF79720	Acf79720 Mosquito
33	73	5.6	6292	4	AAS46735	Aas46735 Tumour su
34	69.6	5.3	7892	6	ABK40056	Abk40056 Human che
35	68.6	5.3	8136	6	ABK39957	Abk39957 Human che
36	68.6	5.3	8136	6	ABL32555	Abl32555 Human imm
37	68.4	5.2	11836	4	AAS45395	Aas45395 Chemical1
38	68.4	5.2	11836	6	ABK28240	Abk28240 DNA trans
39	68.2	5.2	8392	6	ABL33490	Abl33490 Human imm
40	68	5.2	12405	4	AAS45330	Aas45330 Chemical1
41	68	5.2	12405	6	ABK28169	Abk28169 DNA trans
42	68	5.2	12405	6	AAS61143	Aas61143 Human gen
43	67	5.1	8076	6	ABK39955	Abk39955 Human che
44	66.6	5.1	8056	7	ABZ10246	Abz10246 Haematopo
45	66.4	5.1	6669	6	ABL32193	Abl32193 Human imm

ALIGNMENTS

```
RESULT 1
AAQ77857
ID   AAQ77857 standard; DNA; 2346 BP.
XX
AC   AAQ77857;
XX
DT   25-MAR-2003  (revised)
DT   23-JUN-1995  (first entry)
XX
DE   Mycoplasma gallisepticum antigen TM-67 coding sequence.
XX
KW   recombinant avipox virus; live vaccine; mycoplasma antigen; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             202..2049
FT                   /tag= a
FT                   /product= "Mycoplasma_antigen"
FT   mat_peptide     202..2046
FT                   /tag= b
FT                   /codon= seq:TGA,aa:Trp
XX
PN   WO9423019-A1.
XX
PD   13-OCT-1994.
XX
PF   31-MAR-1994;   94WO-JP000541.
XX
PR   31-MAR-1993;   93JP-0074139.
PR   30-SEP-1993;   93JP-0245625.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Saito S,  Ohkawa S,  Saeki S,  Ohsawa I,  Funato H,  Iritani Y;
PI   Aoyama S,  Takahashi K;
XX
DR   WPI; 1994-333181/41.
DR   P-PSDB; AAR63230.
XX
PT   Recombinant avipox virus combining DNA encoding a polypeptide -
PT   exhibiting antigenicity of mycoplasma, useful for the production of a
PT   live vaccine.
XX
PS   Claim 4; Page 87-91; 123pp; Japanese.
```

```
XX
CC   A restriction fragment of the insert of M.gallisepticum genomic clone pUM
CC   -67 containing an open reading frame was sequenced (AAQ77857). The ORF
CC   encodes an antigenic polypeptide. A recombinant avipox virus comprising
CC   the coding sequence can be used as a live vaccine to protect against
CC   infection by Mycoplasma gallisepticum. (Updated on 25-MAR-2003 to correct
XX   PN field.)
SQ   Sequence 2346 BP; 863 A; 385 C; 360 G; 731 T; 0 U; 7 Other;

Query Match             98.2%;  Score 1282;  DB 2;  Length 2346;
Best Local Similarity   98.9%;  Pred. No. 9.8e-199;
Matches 1291;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          1 AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAAAACACAAAATCTTCTAACAA 60

Qy         61 AATCCTAAATAAATAAGCCGTAAATTAACTAACAATTAAAAATTAAAAAAATGGTTTTTCTTATC 120
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db         61 AATCCTAAATAAATAAGCCGTAAATTAACTAACAATTAAAAAATTAAAAAAATGGTTTTTCTTATC 120

Qy        121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTATTTTTAGTCATCTTTTAAGAT 180
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTATTTTTTAGTCATCTTTTAAGAT 180

Qy        181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy        241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300

Qy        301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGGAGTTAACAGAT 360
              |||| |||||||||||||||||||||||||||||| ||||||||||||||||||||| |||
Db        301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGGAATGGAGTTAACTGAT 360

Qy        361 CTAATCAATGCTAAAGCGATGACATGCAAGACATGCTCACTACAGACTATGCTGAAGCT 420
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        361 CTAATCAATGCTAAAGCAAGACATGCAAGACATGCTCACTACAAGACTATGCTAAGCT 420

Qy        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy        481 GAACAACTAAAAAATGGCTAAAAACTAATTAGAATCAGCCATCAACCAAGCTAAATCACGGAT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAACAACTAAAAAATGGCTAAAAACTAATTAGAATCAGCCATCAACCAAGCTAAATCACGGAT 540

Qy        541 AAAAACGACTTTTGATATGAACACCCCAAATTTAGTGAAGCATACAAAGCACTAAAAACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAAACGACTTTTGATATGAACATCCCAAATTTAGTGAAGCATACAAAGCACTAAAAACC 600

Qy        601 ACTTTAGAACAACGTGCTACTACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTTTAGAACAACGTGCTACTACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT 660

Qy        661 CGCAATAATTTAGTGGATCTATACAAATAAAGTAGTAGTTTAATAACTAAAACACTAGAT 720
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTAATAATTTAGTGGATCTATACAAATAAATGTAGTAGTTTAATAACTAAAACACTAGT 720

Qy        721 CCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTTGTGGAATATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCACTAAATGGGGGAATGCTTTTAGATTCTAATGAGATTACTACAGTTTGTGGAATATT 780

Qy        781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTAATATTATCTAATAGT 840
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATAATAACGTTATCAACTATTAATGAACAAAAGACTAATGCTAATATTATCTAATAGT 840

Qy        841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTGTAGGGACTTTTACAAACGCT 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTGTAGGGACTTTTACAAACGCT 900
```

```
Qy        901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACACCGTCAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACACCGTCAAT 960

Qy        961 TATAAATATGCAAGAAGGAGACCGTTGGAATGGTGATGAACCTTCAAGTGAACCTAGAATTCTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATAAATATGCAAGAAGGAGACCGTTGGAATGGTGATGAACCTTCAAGTGAACCTAGAATTCTTGCA 1020

Qy       1021 AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTAGCTGGAACAACAACACGAAG 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTAGCTGGAACAACAACACGAAG 1080

Qy       1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140

Qy       1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy       1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260

Qy       1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT 1306
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT 1306
```

```
RESULT 2
AAQ77854
ID   AAQ77854 standard; DNA; 1387 BP.
XX
AC   AAQ77854;
XX
DT   25-MAR-2003 (revised)
DT   23-JUN-1995 (first entry)
XX
DE   Mycoplasma gallisepticum 40kD antigen coding sequence.
XX
KW   recombinant avipox virus; live vaccine; mycoplasma 40kD antigen; TTM-1;
KW   ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             202..1308
FT                   /tag= a
FT                   /product= "40kD antigen"
FT                   /note= "ORF includes 2 NNN codons"
XX
PN   WO9343019-A1.
XX
PD   13-OCT-1994.
XX
PF   31-MAR-1994;  94WO-JP000541.
XX
PR   31-MAR-1993;  93JP-00074139.
PR   30-SEP-1993;  93JP-00245625.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Saito S, Ohkawa S, Saeki S, Ohsawa I, Funato H, Iritani Y;
PI   Aoyama S, Takahashi K;
XX
DR   WPI; 1994-333181/41.
DR   P-PSDB; AAR63227.
XX
PT   Recombinant avipox virus combining DNA encoding a polypeptide -
PT   exhibiting antigenicity of mycoplasma, useful for the production of a
PT   live vaccine.
```

```
XX
PS   Claim 4; Page 71-74; 123pp; Japanese.
XX
CC   The plasmid pUTTM-1P contains a sequence (the TTM-1 gene) coding for the
CC   40kD antigen of Mycoplasma gallisepticum under the control of a synthetic
CC   promoter. A 1300 bp restriction fragment containing the promoter-ORF
CC   sequence was excised and was used in the construction of plasmid pNZ7929-
CC   R2. This in turn was involved in the construction of a recombinant avipox
CC   virus vector comprising the TTM-1 gene, DNA encoding the signal membrane
CC   anchor peptide from Newcastle Disease Virus haemagglutinin neuraminidase
CC   and FPV sequences. The recombinant avipox virus is useful as a live
CC   vaccine to protect against infection by Mycoplasma gallisepticum.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 1387 BP; 531 A; 220 C; 203 G; 427 T; 0 U; 6 Other;

Query Match             96.8%;  Score 1264;  DB 2;  Length 1387;
Best Local Similarity   98.3%;  Pred. No. 8.2e-196;
Matches 1284;  Conservative    0;  Mismatches   21;  Indels    1;  Gaps    1;

Qy          1 AAAAACATCAGATTGTTAATCTGATATCTTGCTAAAAAAACACAAAATCTTCTAACAA   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAACATCAGATTGTTAATCTGATATCTTGCTAAAAAAACACAAAATCTTCTAACAA   60

Qy         61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAATGGTTTTCTTATC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAATGGTTTTCTTATC  120

Qy        121 AACCAAAATTCTCTAGTAATAAACGCTTATTATTTTTATTTTTAGTCATCTTTTAAGAT  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACCAAAATTCTCTAGTAATAAACGCTTATTATTTTTATTTTTAGTCATCTTTTAAGAT  180

Qy        181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG  240

Qy        241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA  300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA  300

Qy        301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGCGAATGGAGTTAACAGAT  360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGCGAATGGAGTTAACTGAT  360

Qy        361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT  420

Qy        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA  480

Qy        481 GAACAACTAAAAATGGCTAAAAACTAATTGAATCAGCCATCAACCAAGCTAATACCGGAT  540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAACAACTAAAAATGGCTAAAAACTAATTGAATCAGCCATCAACCAAGCTAATACCGGAT  540

Qy        541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATAACAAAGCACTAAAAAACC  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAACGACTTTTGATAATGAACATCCAAATTAGTTGAAGCATAACAAAGCACTAAAAAACC  600

Qy        601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCAACTGCTTATAATCAGATT  660

Qy        661 CGCAATAATTAGTTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAAAACACTAGAT  720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTAATAATTAGTTGGATCTATACAATAATGCTAGTAGTTTAATAACTAAAAAACACTAGAT  720

Qy        721 CCACTCAAATGGGGGAACGCGTTTAGATTCTAAATCGGAATATT  780
              |||||||||||||||||||||||||||||||||||||||||||
Db        721 CCACTCAAATGGGGGAATGCGTTTAGATTCTAAATCGGAATATT  780

Qy        781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840

Qy        841 TTTATTAAAAAGTGATTCAAATTCAAAATAATGAACAACAAGTTTGTGAGGGACTTTTACAAACGCT  900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTTATTAAAAAGTGATTCAAATTCAAAATAATGAACAACAAGTTTGTGAGGGACTTTTACAAACGCT  900

Qy        901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACTAACCCCGTCAAT  960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACTAACCCCGTCAAT  960

Qy        961 TATAAATATGCAAGAGGACGCCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        961 TATAAATATGCAAGAGGACGCCGTTNNNAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020

Qy       1021 AACACGAATAGTATCACAGATGTTTCTTGATTATAGTTTAGTGGAACAAACACGAAG 1080
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1021 AACACGAATAGTATCACAGATGTTTCTNNNNATTTATAGTTTAGTGGAACAAACACGAAG 1080

Qy       1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATATTTCCCTTATAAG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATATTTCCCTTATAAG 1140

Qy       1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAATAATGGAAAT 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAATAATGGAAAT 1200

Qy       1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACTACAGCTAATCCAACT 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACTACAGCTAATCCAACT 1260

Qy       1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCGAGTT 1306
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCGAGTT 1305


RESULT 3
AAQ53419
ID   AAQ53419 standard; DNA; 1387 BP.
XX
AC   AAQ53419;
XX
DT   25-MAR-2003  (revised)
DT   16-JUN-1994  (first entry)
XX
DE   Mycoplasma gallisepticum antigen coding sequence.
XX
KW   Vaccine; mycoplasma infection; poultry; fowl; ds.
OS   Mycoplasma gallisepticum.
XX
FH   Key            Location/Qualifiers
FT   misc_difference 51
FT                  /*tag= b
FT                  /note= "only partially legible in specification; either C
FT                  or G"
FT   misc_difference 54
FT                  /*tag= c
FT                  /note= "only partially legible in specification; either C
FT                  or G"
FT   misc_difference 58
FT                  /*tag= d
FT                  /note= "only partially legible in specification; either C
FT                  or G"
FT   misc_difference 64
FT                  /*tag= e
FT                  /note= "only partially legible in specification; either C
FT                  or G"
FT   CDS            202..1308
FT                  /*tag= a
FT                  /product= "40kD antigen"
FT                  /note= "ORF includes 2 NNN codons"
XX
```

```
PN   WO9324646-A1.
XX
PD   09-DEC-1993.
XX
PF   28-MAY-1993;   93WO-JP000715.
XX
PR   29-MAY-1992;   92JP-00138819.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Saito S,  Ohkawa S,  Fujisawa A,  Iritani Y,  Aoyama S;
XX
DR   WPI; 1993-405837/50.
DR   P-PSDB; AAR44493.
XX
PT   Mycoplasma gallisepticum antigen and DNA coding for it - useful for
PT   vaccination of fowl against mycoplasma infections.
XX
PS   Claim 2; Page 23-26; 37pp; Japanese.
XX
CC   The sequence coding for the 40kDa antigen was obtained by PCR
CC   amplification of M.gallisepticum genomic DNA. The antigen reacts with
CC   Mycoplasma-immune or Mycoplasma-infected serum and can be used as a
CC   vaccine to protect fowl from M.gallisepticum infection. (Updated on 25-
CC   MAR-2003 to correct PN field.)
XX
SQ   Sequence 1387 BP; 531 A; 216 C; 203 G; 427 T; 0 U; 10 Other;

Query Match             96.7%;  Score 1262.4;  DB 2;  Length 1387;
Best Local Similarity   98.0%;  Pred. No. 1.5e-195;
Matches 1280;  Conservative    4;  Mismatches   21;  Indels    1;  Gaps   1;

Qy         1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAACACAAAATCTTCTAACAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||| ::||:||:
Db         1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAACACAAAATSTSTAASAA 60

Qy        61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATTGGTTTTTTCTTATC 120
             |||::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AATSCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATTGGTTTTTTCTTATC 120

Qy       121 AACCAAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAGAT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AACCAAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAGAT 180

Qy       181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy       241 TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGCTGTATTGTCATTACTAAAAAA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TTAGGTACAACATCCTTCTTTAGCATTGGGATTTCTAGCTGTATGCTGTATTGTCATTACTAAAAAA 300

Qy       301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGGAGTTAACAGAT 360
             ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GACGCAAACCCAAATAATGGCCAAACCCAATTACACAGCAGCGCGCAATGGAGTTAACTGAT 360

Qy       361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGAGATGCCAAGATTGAAGCT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAAGAGATGCCAAGATTGAAGCT 420

Qy       421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTGAAACAGTTAACATAACCTTAATGCAACATTA 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACATAACCTTAATGCAACACTA 480

Qy       481 GAACAACTAAAAAATGCTAAAACTAATTTAGAATCAGCGCCATCAACCAGCTAATACGGAT 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GACGCAACTAAAAAATGCTAAAACTAATTTAGAATCAGCGCCATCAACCAGCTAATACGGAT 540

Qy       541 AAAACGACTTTTGATAATGAACACCCCAAATTTAGTTGAGCATACAAAAGCACTAAAAACC 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAGCATACAAAAGCACTAAAAACC 600

Qy       601 ACTTTAGAACAACGTGCTACTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ACTTTAGACACAACGTGCTACTACTAACCTTGAAGGTTTCAACTGCTTATAATCAGATT 660

Qy       661 CGCAATAATTAGTGATGTCATACATAAAGCTAGTAGTTAATAACTAAAACACTAGAT 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CGTAATAATTAGTGATCATCATACATAAACAATAGTTAATAACTAAAACACTAGAT 720

Qy       721 CCACTAAATGGGGAACGCTTCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCACTAAATGGGGAATGCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy       781 AATAATACGTTATCAACTAATGAACAAAGACTAAATGCTGATGCATTATCTAATAGT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 AATAATACGTTATCATCAACTATTAATGAACAAAGACTAAATGCTGATGCATTATCTAATAGT 840

Qy       841 TTTATTAAAAAAGTGATTCAACTACAGTTTTGTTCTTTTAGTGCTGATGTAACCCGTCAAT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TTTATTAAAAAAGTGATTCAACTACAGTTTTGTTCTTTTAGTGCTGATGTAACACCCGTCAAT 900

Qy       901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTAGTGCTTTTAGTGCTGATGATTAACCGGTCGA 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTAGTGCTTTTAGTGCTGATGATGAACCTTCAAGTTAGAATTCTTGCA 960

Qy       961 TATAAATATGCAAGAGGACCGTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
             ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db       961 TATAAATATGCAAGAGGACCGTTNNNAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020

Qy      1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAACACGAAG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AACACGAATAGTATCACAGATGTTTCTNNNATTTATAGTTTAGCTGGAACAAACACGAAG 1080

Qy      1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATTCCCTTATAACAA 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATTTATTCCCTTATAAG 1140

Qy      1141 TTGGTTAAAGCAGCTGATGCTAATACGTTGGATTACAATCAAATTAAATAATGGAAAT 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 TTGGTTAAAGCAGCTGATGCTAATACGTTGGATTACAATCAAATTAAATAATGGAAAT 1200

Qy      1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTCAATAATATACTACAGCTAATCCAACT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTCAATAATATACTACAGCTAATCCAACT 1260

Qy      1261 CCAGCAGTTGATGAGATTAAAGTGCTAAAATCGTTTTATCAGGTT 1306
             ||||||||||||||||||||||||||||||||||||||||||||
Db      1261 -CAGCAGTTGATGAGATTAAAGTGCTAAAATCGTTTTATCAGGTT 1305


RESULT 4
AAQ94711
ID   AAQ94711 standard; DNA; 1387 BP.
XX
AC   AAQ94711;
XX
DT   25-MAR-2003 (revised)
DT   12-MAR-1996 (first entry)
XX
DE   Mycoplasma gallisepticum antigenic protein TTM-1 gene.
XX
KW   Antigenic protein; vaccine; poultry; diagnosis; TTM-1; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             202..1308
FT                   /*tag= a
XX
PN   JP07133295-A.
PD   23-MAY-1995.
XX
PF   27-AUG-1993;  93JP-00213102.
```

```
XX
PR   27-AUG-1993;  93JP-00213102.
XX
PA   (JAPG )  NIPPON ZEON KK.
PA   (SHIO )  SHIONOGI & CO LTD.
XX
DR   WPI; 1995-220782/29.
DR   P-PSDB; AAR76955.
XX
PT   A new antigenic protein which reacts with Mycoplasma gallisepticum - is
PT   useful in a component vaccine for use against poultry infected with M.
PT   gallisepticum.
XX
PS   Claim 6; Fig 5-6; 33pp; Japanese.
XX
CC   AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein TTM-
CC   1. TTM-1 can be used as a vaccine for M. gallisepticum infectious
CC   diseases in poultry, and as a diagnostic agent for M. gallisepticum.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1387 BP; 531 A; 222 C; 427 T; 0 U; 6 Other;

Query Match             96.5%;  Score 1260.8;  DB 2;  Length 1387;
Best Local Similarity   98.2%;  Pred. No. 2.7e-195;
Matches 1282;  Conservative    0;  Mismatches  23;  Indels    1;  Gaps    1;

Qy          1 AAAAACAATCAGATTGTTAATCTGATAATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
              |||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAACATCAGATTGTTAATCTGATAATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy         61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAAATTAAAAAAAATGGTTTTTTCTTATC 120
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAAATTAAAAAAAATGGTTTTTTCTTATC 120

Qy        121 AACCAAAATTCTCTAGTAATAACGCTTATTATTTATTTTTATTTTAGTCATCTTTTAAGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACCAAAATTCTCTAGTAATAACGCTTATTATTTATTTTTATTTTAGTCATCTTTTAAGAT 180

Qy        181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy        241 TTAGGTACAACATCCTTTCTTCTTAGCATTGGATTCTCAGCTGTATGTCTGTATTACTAAAAAA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTAGGTACAACATCCTTTCTTCTTAGCATTGGATTCTCAGCTGTATGTCTGTATTACTAAAAAA 300

Qy        301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAGGCAGCGAGGCGGCGAATGGAGTTAACAGAT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACGCAAACCCAAATAATGGCCAAACCCAATTAGAGGCAGCGAGGCGGCGAATGGAGTTAACTGAT 360

Qy        361 CTAATCAATGCTAAAGCGATGACATTAGTTCACTACAAGACTATGCCAAGATTGAAGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAATCAATGCTAAAGCGATGACATTAGTTCACTACAAGACTATGCCAAGATTGAAGCT 420

Qy        421 AGTTTATCATCTGCTTATAGTGCTAAAACAGTTAACAATAACCTTAATGCAACATTA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTTTATCATCTGCTTATAGTGCTAAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy        481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACAAGCTAATACGGAT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACAAGCTAATACGGAT 540

Qy        541 AAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATAACAAGCACTAAAAACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAACGACTTTTGATAATGAACATCCAAATTTAGTTCAACCATACAAAGCACTAAAAACC 600

Qy        601 ACTTTAGACAACACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAATCAAATT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTTTAGACAACACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAATCAGATT 660

Qy        661 CGCAATAATTTAGTGGATCTAGTCACTACAATAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTAATAATTTAGTGGATCTAGTCACTACAATAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
```

```
Qy        721 CCACTAAATGGGGACGCTTTAGATTCTAATGAAGTTAATCACAGTAGTTAATCGGAATATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCACTAAATGGGGACGGATGCTTTAGATTCTAATGAAGTTAATCACAGTAGTTAATCGGAATATT 780

Qy        781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT 840

Qy        841 TTTATTAAAAAGTGATTCAAAATAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTTATTAAAAAGTGATTCAAAATAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy        901 AATGTTCAACCTTCAAACTACAGTTTGTTGTTGCTTTGTGCTGTGATGTAACACCCGTCAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTTCAACCTTCAAACTACAGTTTGTTGTTGCTTTGTGCTGTGATGTAACACCCGTCAAT 960

Qy        961 TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
              |||||||||||||||||||||||||||||||    |||||||||||||||||||||||
Db        961 TATAAATATGCAAGAAGGACCGTTNNNAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020

Qy       1021 AACACGAATAGTATCACAGATGTTTCTTGGATTATAGTTTAGCTGGAACAAACACGAAG 1080
              ||||||||||||||||||||||||||||||     |||||||||||||||||||||||
Db       1021 AACACGAATAGTATCACAGATGTTTCTNNNNATTATAGTTTAGCTGGAACAAACACGAAG 1080

Qy       1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCCAACTGGTTATTATATTTCCCTTATAAG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCCAACTGGTTATTATATTTCCCTTATAAG 1140

Qy       1141 TTGGTTAAAGCAGCTGATGCGTAATACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGGTTAAAGCAGCTGATGCGTAATACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy       1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260

Qy       1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
              |||||||||||||||||||||||||||||| ||||||||||||||
Db       1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1305
```

```
RESULT 5
AAQ77856
ID   AAQ77856 standard; DNA; 1935 BP.
XX
AC   AAQ77856;
XX
DT   25-MAR-2003 (revised)
DT   23-JUN-1995 (first entry)
XX
DE   Mycoplasma gallisepticum antigen TM-66 coding sequence.
XX
KW   recombinant avipox virus; live vaccine; mycoplasma antigen; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..1886
FT                   /*tag= a
FT                   /product= "Mycoplasma_antigen"
FT   mat_peptide     54..1883
FT                   /*tag= b
FT                   /codon= seq:TGA,aa:Trp
XX
PN   WO9423019-A1.
XX
PD   13-OCT-1994.
XX
PF   31-MAR-1994; 94WO-JP000541.
XX
PR   31-MAR-1993; 93JP-00074139.
PR   30-SEP-1993; 93JP-00245625.
```

```
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Saito S, Ohkawa S, Saeki S, Ohsawa I, Funato H, Iritani Y;
PI   Aoyama S, Takahashi K;
XX
DR   WPI; 1994-333181/41.
DR   P-PSDB; AAR63229.
XX
PT   Recombinant avipox virus combining DNA encoding a polypeptide -
PT   exhibiting antigenicity of mycoplasma, useful for the production of a
PT   live vaccine.
XX
PS   Claim 4; Page 78-81; 123pp; Japanese.
XX
CC   A restriction fragment of the insert of M.gallisepticum genomic clone pUM
CC   -66 containing an open reading frame was sequenced (AAQ77856). The ORF
CC   encodes an antigenic polypeptide. A recombinant avipox virus comprising
CC   the coding sequence can be used as a live vaccine to protect against
CC   infection by Mycoplasma gallisepticum. (Updated on 25-MAR-2003 to correct
CC   PN field.)
XX
SQ   Sequence 1935 BP; 728 A; 314 C; 307 G; 586 T; 0 U; 0 Other;

Query Match             83.2%;  Score 1086.8;  DB 2;  Length 1935;
Best Local Similarity   96.7%;  Pred. No. 3.8e-167;
Matches 1120;  Conservative    0;  Mismatches  37;  Indels    1;  Gaps    1;

Qy    150 TTTATTTTTATTTTTAGT-CATCTTTAAGATATAAATATATCTTAAATATTCTATGAATA 208
          ||||||||||||| ||||| |||||||||||||||| ||||||||||||||||||||||
Db      1 TTTATTTTTATTTTTGGTAAATCTTTAAAATATAAATATATTTAATATTCTATGAATA 60

Qy    209 AGAAAAGAATCATCTTAAAGACTATTAGTTGTTAGGTACAACATCCTTTCTTAGCATTG 268
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     61 AAAAAGAATCATCTTAAAGACTATTAGCTTGTTAGGTACAACATCCTTTCTTAGTATTG 120

Qy    269 GGATTTCTAGCTGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCC 328
          |||||||||||||||||||||||||        |||||||||||||||||||||||
Db    121 GGATTTCTAGCTGTCTATTACTAATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCC 180

Qy    329 AATTAGAAGCAGGCGAATGGCGAATGGAGTTAACAGATCTAATCAATGCTAAGCGATGACATTAG 388
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATTAGAAGCAGGCGAATGGCGAATGGAGTTAACAGATCTAATCAATGCTAAGCGATGACATTAG 240

Qy    389 CTTCACTACAAGACTATGGCAAGATTGAAGCTAGTTGAGCTTATCATCTGCTTATAGTGAAGCTG 448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCACTACAAGACTATGGCAAGATTGAAGCTAGTTGAGCTTATCATCTGCTTATAGTGAAGCTG 300

Qy    449 AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATT 508
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATT 360

Qy    509 TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAA 568
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAA 420

Qy    569 ATTTAGTTGAAGCATACAAAGCACTAAAACCACTTTAGAACAACGTGCTACTAACCTTG 628
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATTTAGTTGAAGCATACAAAGCACTAAAACCACTTTAGAACAACGTGCTACTAACCTTG 480

Qy    629 AAGGTTTGTCATCAACTGCTTATATATCAATTCGCAATAATTCGGCAATAATTAGT 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAGGTTTGTCATCAACTGCTTATATATCAATTCGCAATAATTCGGCAATAATTAGT 540

Qy    689 AAGCTAGTAGTTAATAACTAAAACACTAGATCCACTAAAATGGGGGAACGGCTTTAGATT 748
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAGCTAGTAGTTAATAACTAAAACACTAGATCCACTAAAATGGGGGAACGGCTTTAGATT 600

Qy    749 CTAATGAGATTACTACAGCTAATAGAATATTAATTAATAATCGTTATCAACTATTAATGAAC 808
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTAATGAGATTACTACAGCTAATAGAATATTAATTAATAATCGTTATCAACTATTAATGAAC 660
```

```
Qy    809 AAAAGACTAATGCTGATGCAATATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATG 868
          ||||||||||||||||||||       |||||||||       |||||||||||||||
Db    661 AAAAGACTAATGCTGATGCATTAGCTCATTAGTTTATTAATAGTTTATTAAAGAAGTGATTCAAAATAATA 720

Qy    869 AACAAAGTTTGTAGGGACTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTG 928
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    721 AACAAAGTTTGTAGGGATGTTTACAAACATAATGTTCAACCTTCAAACTATAGTTTTG 780

Qy    929 TTGCTTTTAGTGCTGATGTAACACCCGTCATTATAAATATGCAAGAAGGACCGTTTGGA 988
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTGCTTTTAGTGCTGATGTAACACCCGTTATTATTATAAATATGCAAGAAGAAGACGGTTTGAA 840

Qy    989 ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTT 1048
          |||||||||||||||||||||||||||||||||       ||||||||||||||||||
Db    841 ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACAACACCAACAGTATTACTGATGTTTCAT 900

Qy   1049 GGATTTATAGTTAGTTGGACAACAAACACGAGTACCAATTTAGTTTTAGCAACTATGGTC 1108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAATTTATAGTTTATCTGGACAACAAACACGAGTACCAATTTAGTTTTTAGCAACTACGGTC 960

Qy   1109 CATCAACTGGTTATTATATTTCCCTATAGTTGGTTAAAGCAGCTGATGCTAATAACG 1168
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CATCAACTGGTTATTATTATTTCCCTTATAGTTGGTTAAAGCGGCTGATGCTAGTAATG 1020

Qy   1169 TTGGATTACAATACAAATTAAATAATGGAAATGTTCAACAAGTTGAGTTGTTGCCACTTCAA 1228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTGGATTACAATACAAACTAAATAATGGAAATGTTCAACAGTTGAGTTTGCCACTTCAA 1080

Qy   1229 CTAGTGCAAATAATACACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA 1288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTAGCGCAAATAATACACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA 1140

Qy   1289 AAATCGTTTTATCAGGTT 1306
          ||||||||||||||||||
Db   1141 AAATCGTTTTATCAGGTT 1158
```

```
RESULT 6
ACF03364
ID   ACF03364 standard; DNA; 1152 BP.
XX
AC   ACF03364;
XX
DT   11-SEP-2003  (first entry)
XX
DE   Mycoplasma gallisepticum TTM-1 gene SEQ ID NO:1.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP1275716-A2.
XX
PD   15-JAN-2003.
XX
PF   11-JUL-2002; 2002EP-00254879.
XX
PR   11-JUL-2001; 2001US-00901572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG ) ZEON CORP.
XX
PI   Okuda T, Saito S, Dorsey KM, Tsuzaki Y;
XX
DR   WPI; 2003-373746/36.
XX
PT   DNA molecule derived from a prokaryotic cell, useful for producing a
PT   vaccine for treating viral infections comprises at least one modified DNA
PT   regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS   Claim 9; Page 30; 70pp; English.
XX
```

CC The present invention describes a DNA molecule derived from a prokaryotic
CC cell, where at least one of the DNA regions encoding NXB (where N =
CC asparagine, X = any amino acid other than proline, and B = serine or
CC threonine) has been modified so that no N-glycosylation occurs during the
CC expression in a eukaryotic cell. Also described: (1) a fused DNA
CC molecule, where a DNA encoding a signal sequence has been ligated to the
CC N-terminal end of the modified DNA molecule as described above so that it
CC may be expressed as a fusion protein; (2) a recombinant virus integrated
CC with the DNA molecule or the fused DNA molecule described above; (3)
CC producing a modified or fusion protein by using the recombinant virus
CC described above, to express a protein encoded by the modified DNA
CC molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC vaccine comprising the recombinant virus. The DNA molecule has virucide
CC and immunostimulant activities. The DNA molecule is useful for producing
CC a vaccine for treating viral infections. The present sequence is used in
CC the exemplification of the present invention
XX
SQ Sequence 1152 BP; 437 A; 190 C; 188 G; 337 T; 0 U; 0 Other;

Query Match   78.7%; Score 1028; DB 7; Length 1152;
Best Local Similarity   100.0%; Pred. No. 1.2e-157;
Matches 1028; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy  279 CTGTATGTCTATTACTAAAAAAAGATGCAAACCCAAATAATGCCAAACCCAATTAGAAGC 338
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6 CTGTATGTCTATTACTAAAAAAAGATGCAAACCCAAATAATGCCAAACCCAATTAGAAGC 65

Qy  339 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACA 398
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AGCGCGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACA 125

Qy  399 AGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAA 458
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126 AGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAA 185

Qy  459 CAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGC 518
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  186 CAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGC 245

Qy  519 CATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGA 578
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  246 CATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGA 305

Qy  579 AGCATACAAAGCACTAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGTC 638
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  306 AGCATACAAAGCACTAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGTC 365

Qy  639 ATCAACTGCTTATAATCGAATTCGCAATAATTCAAATTGGATCATACAATAAAGCTAGTAG 698
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  366 ATCAACTGCTTATAATCGAATTCGCAATAATTCAAATTGGATCATACAATAAAGCTAGTAG 425

Qy  699 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTAGATTCTAATGAGAT 758
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  426 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTAGATTCTAATGAGAT 485

Qy  759 TACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 818
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  486 TACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 545

Qy  819 TGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTT 878
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  546 TGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTT 605

Qy  879 TGTAGGGACTTTACAAACGCTAATGTCAACCTTCAAACTACAGTTTTGTTGCTTTTAG 938
        ||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db  606 TGTAGGGACTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAG 665

Qy  939 TGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGA 998
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  666 TGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGA 725

Qy  999 ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAG 1058
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  726 ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAG 785

Qy 1059 TTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGG 1118
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  786 TTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGG 845

Qy 1119 TTATTTATATTCCCTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACA 1178
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  846 TTATTTATATTCCCTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACA 905

Qy 1179 ATACAAATTAAATAATGGAAATGTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA 1238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  906 ATACAAATTAAATAATGGAAATGTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA 965

Qy 1239 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGATGAGATTAAAGTTGCTAAAATCGTTTT 1298
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  966 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGATGAGATTAAAGTTGCAAAATCGTTTT 1025

Qy 1299 ATCAGGTT 1306
        ||||||||
Db 1026 ATCAGGTT 1033
```

RESULT 7
AAT96595
ID   AAT96595 standard; DNA; 1371 BP.
XX
AC   AAT96595;
XX
DT   15-JUL-1998 (first entry)
XX
DE   Hybrid Marek's disease virus/M. gallisepticum gene 40 K-S.
XX
KW   Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW   antigen; vaccine; poultry; ds.
XX
OS   Marek's disease gammaherpesvirus.
OS   Mycoplasma gallisepticum.
OS   Chimeric.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1371
FT   /tag= a
FT   /product= "protein 40 K-S"
FT   misc_feature   1..192
FT   /tag= b
FT   /note= "derived from Marek's disease virus gB gene"
FT   misc_feature   193..1368
FT   /tag= c
FT   /note= "derived from gene encoding M. gallisepticum
FT   antigen"
XX
PN   WO9736924-A1.
XX
PD   09-OCT-1997.
XX
PF   28-MAR-1997;   97WO-JP001084.
XX
PR   29-MAR-1996;   96JP-00103548.
XX
PA   (JAPG) NIPPON ZEON KK.
XX
PI   Saito S, Tsuzaki Y, Yanagida N;
XX
DR   WPI; 1997-503046/46.
DR   P-PSDB; AAW36050.
XX
PT   Fusion protein comprising herpes virus outer membrane protein and
PT   antigenic polypeptide - for prevention of infection by Mycoplasma
PT   gallisepticum, especially in poultry.
XX
PS   Disclosure; Page 16-19; 51pp; Japanese.
XX

CC   This sequence encodes the chimeric protein 40 K-S which comprises a
CC   fragment of the Marek's disease virus outer membrane protein gB fused to
CC   an antigenic protein from Mycoplasma gallisepticum. The chimeric protein
CC   can be used in recombinant live vaccines for prevention of infection by
CC   Mycoplasma gallisepticum, especially as the outer membrane protein shows
CC   antigenicity in poultry
XX
SQ   Sequence 1371 BP; 491 A; 242 C; 232 G; 406 T; 0 U; 0 Other;

Query Match   78.4%; Score 1024.2; DB 2; Length 1371;
Best Local Similarity   99.7%; Pred. No. 5.1e-157;
Matches 1026; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   278 GCTGTATGTCTATTACTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG 337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 GCTGTATGTCTATTACTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG 247

Qy   338 CAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTAC 397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 CAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTAC 307

Qy   398 AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATGAGCTGAAACAGTTA 457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATGAGCTGAAACAGTTA 367

Qy   458 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTAGAATCAG 517
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTAGAATCAG 427

Qy   518 CCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACACCCAAATTTAGTTG 577
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACACCCAAATTTAGTTG 487

Qy   578 AAGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGT 637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 AAGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGT 547

Qy   638 CATCAACTGCTTATATCAATTCGCAATAATTAGTGATCTATACAATAAAGCTAGTA 697
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 CATCAACTGCTTATATCAATTCGCAATAATTAGTGATCTATACAATAAAGCTAGTA 607

Qy   698 GTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGA 757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 GTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGA 667

Qy   758 TTACTACAGTTAATCGGAATTATAATTAATAATACGTTATCAACTATTAATGAACAAAGACTA 817
          ||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||||
Db   668 TTACTACAGTTAATAGGAATTATAATTAATAATACGTTATCAACTATTAATGAACAAAGACTA 727

Qy   818 ATGCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTT 877
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 ATGCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTT 787

Qy   878 TTGTAGGGACTTTACAAACGCTAATGTTCAAACTACAGTTTTGTTGCTTTTA 937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 TTGTAGGGACTTTACAAACGCTAATGTTCAAACTACAGTTTTGTTGCTTTTA 847

Qy   938 GTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCCGTTGGAATGGTGATG 997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 GTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCCGTTGGAATGGTGATG 907

Qy   998 AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATA 1057
          ||||||||||||||||||||||||||| ||||||||||||| |||||| ||||||||||
Db   908 AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATA 967

Qy   1058 GTTTAGCTGAACAAACACCGAAGTACCGAATTAGTTTAGCAACTATGTTCCATCAACTG 1117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GTTTAGCTGGAACAAACACCGAAGTACCAAATTAGTTTAGCAACTATGTTCCATCAACTG 1027

Qy   1118 GTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTCAATAACGTTGGATTAC 1177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 GTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGTCAATAACGTTGGATTAC 1087

Qy   1178 AATACAAATTAAATAATAATGGAAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAA 1237
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 AATACAAATTAAATAATAATGGAAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAA 1147

Qy   1238 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1297
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1207

Qy   1298 TATCAGTT 1306
          ||||||||
Db   1208 TATCAGTT 1216

RESULT 8
AAT96596
ID   AAT96596 standard; DNA; 3261 BP.
XX
AC   AAT96596;
XX
DT   15-JUL-1998 (first entry)
XX
DE   Hybrid Marek's disease virus/M. gallisepticum gene 40 K-C.
XX
KW   Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW   antigen; vaccine; poultry; ds.
XX
OS   Marek's disease gammaherpesvirus.
OS   Mycoplasma gallisepticum.
OS   Chimeric.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..3261
FT   /*tag= a
FT   /product= "protein 40 K-C"
FT   misc_feature   1..2016
FT   /*tag= b
FT   /note= "derived from Marek's disease virus gB gene"
FT   misc_feature   2077..3258
FT   /*tag= c
FT   /note= "derived from gene encoding M. gallisepticum
FT   antigen"
XX
PN   WO9736924-A1.
XX
PD   09-OCT-1997.
XX
PF   28-MAR-1997; 97WO-JP001084.
XX
PR   29-MAR-1996; 96JP-00103548.
XX
PA   (JAPG) NIPPON ZEON KK.
XX
PI   Saito S, Tsuzaki Y, Yanagida N;
XX
DR   WPI; 1997-503046/46.
DR   P-PSDB; AAW36051.
XX
PT   Fusion protein comprising herpes virus outer membrane protein and
PT   antigenic polypeptide - for prevention of infection by Mycoplasma
PT   gallisepticum, especially in poultry.
XX
PS   Disclosure; Page 22-30; 51pp; Japanese.
XX
CC   This sequence encodes the chimeric protein 40 K-C which comprises a
CC   fragment of the Marek's disease virus outer membrane protein gB fused to
CC   an antigenic protein from Mycoplasma gallisepticum. The chimeric protein
CC   can be used in recombinant live vaccines for prevention of infection by
CC   Mycoplasma gallisepticum, especially as the outer membrane protein shows
CC   antigenicity in poultry
XX
SQ   Sequence 3261 BP; 1099 A; 617 C; 648 G; 897 T; 0 U; 0 Other;

Query Match   78.3%; Score 1022.6; DB 2; Length 3261;

Best Local Similarity 99.6%; Pred. No. 9e-157;
Matches 1025; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy    278  GCTGTATGTCTATTACTAACTAAAAAAGATGCAAACCCAATTAGAAG  337
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2078  GCTGTATGTCTATTACTAACTAAAAAAGATGCAAACCCAATTAGAAG  2137

Qy    338  CAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC  397
           ||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db   2138  CAGCCCGAATGGAGTTAACAGATCTAACAAATGCTAAAGCGATGACATTAGCTTCACTAC  2197

Qy    398  AAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA  457
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2198  AAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA  2257

Qy    458  ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAG  517
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2258  ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAG  2317

Qy    518  CCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG  577
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2318  CCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG  2377

Qy    578  AAGCATACAAAGCACTAAAACCCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT  637
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2378  AAGCATACAAAGCACTAAAACCCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT  2437

Qy    638  CATCAACTGCTTATAATCAAATTCGCAATAATTAGTTGGATCTATACAATAAAGCTAGTA  697
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2438  CATCAACTGCTTATAATCAAATTCGCAATAATTAGTTGGATCTATACAATAAAGCTAGTA  2497

Qy    698  GTTTAATAACTAAAAACACTAGATCCACTAAAAATGGGAGGAACGCTTTAGATTCTAATGAGA  757
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2498  GTTTAATAACTAAAAACACTAGATCCACTAAAAATGGGAGGAACGCTTTAGATTCTAATGAGA  2557

Qy    758  TTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAGACTA  817
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2558  TTACTACAGCTAATAAGAATATTAATAATACGTTATCAACTATTAATGAACAAAGACTA  2617

Qy    818  ATGCTGATGCAATATCTAATAGTTTTATTAAAAAGTGATTCAAATAATGAACAAAGTT  877
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2618  ATGCTGATGCAATATCTAATAGTTTTATTAAAAAGTGATTCAAATAATGAACAAAGTT  2677

Qy    878  TTGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACACAGTTTGTTGCTTTTA  937
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2678  TTGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACACAGTTTGTTGCTTTTA  2737

Qy    938  GTGCTGATGTAACACCCGTCAATTATAAAATATAAAATGCCAAGAAGGACCGTTTGGAATGGTGATG  997
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2738  GTGCTGATGTAACACCCGTCAATTATAAAATATAAAATGCCAAGAAGGACCGTTTGGAATGGTGATG  2797

Qy    998  AACCTTCAAGTAGAATTCTTGCAAACACAGCAGTACCAATTAGTTTTAGCAACTATGGATTATA  1057
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2798  AACCTTCAAGTAGAATTCTTGCAAACACAGCAGTATCCAATTAGTTTTAGCAACTATGGATTATA  2857

Qy   1058  GTTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCACTAATGGTCCATCAACTG  1117
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2858  GTTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCACTAATGGTCCATCAACTG  2917
```

RESULT 9
AAT04076
ID   AAT04076 standard; DNA; 1015 BP.
XX
AC   AAT04076;
XX
DT   19-JUL-1996 (first entry)
XX
DE   M.gallisepticum DNA sequence II encodes 261 amino acid protein.
XX
KW   Detection; probe; primer; PCR; amplification; secretion; lung;
KW   avian chronic respiratory disease; respiratory tract; nasal cavity; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key   Location/Qualifiers
FT   CDS   202..987
FT   /*tag= a
XX
PN   JP07236498-A.
XX
PD   12-SEP-1995.
XX
PF   25-FEB-1994;   94JP-00052764.
PR   25-FEB-1994;   94JP-00052764.
XX
PA   (JAPG) NIPPON ZEON KK.
PA   (SHIO) SHIONOGI & CO LTD.
XX
DR   WPI; 1995-347462/45.
DR   P-PSDB; AAR79911.
XX
PT   Detection of Mycoplasma gallisepticum - for the quick detection, i.e.
PT   within one day, of avian chronic respiratory diseases.
XX
PS   Claim 3; Page 10-11; 11pp; Japanese.
XX
CC   This is the nucleotide sequence of a fragment of the Mycoplasma
CC   gallisepticum genome which codes for a 261 amino acid protein. This
CC   sequence and the sequence of AAT04075 (encoding a 661 amino acid protein)
CC   can be used to detect M.gallisepticum using probes based on nucleotides
CC   1125-1648 and primers based on nucleotides 449-466, the complement of
CC   bases 893-919, 1908-1934 and the complement of bases 2184-2210 of
CC   AAT04075, and a probe based on nucleotides 718-41 of this sequence. The
CC   method using these sequences is faster i.e. is able to detect
CC   M.gallisepticum, which causes avian chronic respiratory diseases, within
CC   one day, from avian secretions, washings from the lung, respiratory
CC   tract, nasal cavity, etc
XX
SQ   Sequence 1015 BP; 400 A; 162 C; 138 G; 315 T; 0 U; 0 Other;

Query Match   75.8%;   Score 989.4;   DB 2;   Length 1015;
Best Local Similarity 98.4%;   Pred. No. 2.2e-151;
Matches 999; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy     1  AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA  60
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAAAACACAAAATCTTCTAACAA  60

Qy    61  AATCCTAAATAAATAAGGCCGTTAAAATAACTAAAAAATTAAAAAAAAAATGGTTTTTCTTATC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AATCCTAAATAAATAAATAAGGCCGTTAAAATAACTAAAAAATTAAAAAAAAAATGGTTTTTCTTATC  120

Qy   121  AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGAT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAGTAAGAT  180

Qy   181  ATAAATAATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1298  TATCAGGTT  1306
           |||||||||
```

```
Db         ATAAATATATCTTAATATTCTATGAATAGAAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy     241 TTAGGTACAACACATCCTTTCTTAGCATGGGATTCTAGCTGTATGTCTATTACTAAAAAA 300
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     241 TTAGGTACAACATCCTTTCTTAGCATGGGATTCTAGCTGTATGTCTATTACTAAAAAA 300

Qy     301 GATGCAAACCCAAATAATATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGAT 360
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAGCAGCGCGAATGGAGTTAACTGAT 360

Qy     361 CTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 420

Qy     421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy     481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540

Qy     541 AAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 600

Qy     601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
           ||||||||||||||| |||||||||||||| |||||||||||||||||
Db     601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCAACTGCTTATAATCAGATT 660

Qy     661 CGCAATAATTTAGTTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
           |||||||||||||||||||| |||||||||||||||||||||
Db     661 CGTAATAATTTAGTTGGATCTATACAATAATGTCTAGTAGTTTAATAACTAAAACACTAGAT 720

Qy     721 CCACTAAATGGGGGAACGCTTTTAGATTCTAAATGAGAATTACTACAGTTAATCGGAATATT 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCACTAAATGGGGGAACGCTTTTAGATTCTAAATGCGAATTAATCGGAATATT 780

Qy     781 AATAATACGTTATCAACTATTAATGAACAAAGACTATGTGATGCATTATCTAATAGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AATAATACGTTATCAACTATTAATGAACAACAAAGACTTTGTGATGCATTATCTAATAGT 840

Qy     841 TTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTGAGGGACTTTACAAACGCT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTGAGGGACTTTACAAACGCT 900

Qy     901 AATGTTCAACCTTCAAACTACAGTTTTGTCTTTTAGTGCTGATGTAACACCCGTCAAT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AATGTTCAACCTTCAAACTACAGTTTTGTCTTTTAGTGCTGATGTAACACCCGTCAAT 960

Qy     961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TATAAATATGCAAGAAGGACCGTTTGAAATGGTGATGAACCTTCAAGTAGAATTC 1015
```

RESULT 10
ACF03382
ID   ACF03382 standard; DNA; 1082 BP.
XX
AC   ACF03382;
XX
DT   11-SEP-2003 (first entry)
XX
DE   M. gallisepticum TTM-1 portion of modified pNZ40K-S gene SEQ ID NO:24.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP127571-A2.
XX
PD   15-JAN-2003.

```
XX   11-JUL-2002; 2002EP-00254879.
PF
XX
PR   11-JUL-2001; 2001US-00901572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG) ZEON CORP.
XX
PI   Okuda T, Saito S, Dorsey KM, Tsuzaki Y;
XX
DR   WPI; 2003-373746/36.
DR   P-PSDB; ABR57374.
XX
PT   DNA molecule derived from a prokaryotic cell, useful for producing a
PT   vaccine for treating viral infections comprises at least one modified DNA
PT   regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS   Example 1; Page 42-43; 70pp; English.
XX
CC   The present invention describes a DNA molecule derived from a prokaryotic
CC   cell, where at least one of the DNA regions encoding NXB (where N =
CC   asparagine, X = any amino acid other than proline, and B = serine or
CC   threonine) has been modified so that no N-glycosylation occurs during the
CC   expression in a eukaryotic cell. Also described: (1) a fused DNA
CC   molecule, where a DNA encoding a signal sequence has been ligated to the
CC   N-terminal end of the modified DNA molecule as described above so that it
CC   may be expressed as a fusion protein; (2) a recombinant virus integrated
CC   with the DNA molecule or the fused DNA molecule described above; (3)
CC   producing a modified or fusion protein by using the recombinant virus
CC   described above, to express a protein encoded by the modified DNA
CC   molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC   vaccine comprising the recombinant virus. The DNA molecule has virucide
CC   and immunostimulant activities. The DNA molecule is useful for producing
CC   a vaccine for treating viral infections. The present sequence is used in
CC   the exemplification of the present invention
XX
SQ   Sequence 1082 BP; 403 A; 184 C; 183 G; 312 T; 0 U; 0 Other;

Query Match       70.6%;   Score 921.6; DB 7;   Length 1082;
Best Local Similarity 98.0%;   Pred. No. 2e-140;
Matches 933;   Conservative 0;   Mismatches 19;   Indels 0;   Gaps 0;

Qy   355 ACAGAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT 414
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     1 ACAGAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT 60

Qy   415 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA 474
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTGCAGGCA 120

Qy   475 ACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT 534
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT 180

Qy   535 ACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 594
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 240

Qy   595 AAAACCACTTTAGAACAACGTGCTACTAACGTGCTACTAACCTTGAAGGTTTCATCAACTGCTAATAAT 654
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAAACCACTTTAGAACAACGTGCTACTAACGTGCTACTAACCTTGAAGGTTTCATCAACTGCTAATAAT 300

Qy   655 CAAATTCGCAATAATTTAGTTGGATCGATCTATACAATAACCTAGTAGTTTAATAACTAAAACA 714
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAAATTCGCAATAATTTAGTTGGATCGATCTATACAATAACCTAGTAGTTTAATAACTAAAACA 360

Qy   715 CTAGATCCACTAAATGGGGGAACGCGTTTTAGATTACTACAGTTAATCGG 774
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTAGATCCACTAAATGGGGGAACGCGTTTTAGATTACTACAGTTAATAAG 420

Qy   775 AATATTAATAATACGTTATCAACTATTAATGAACAAAGACTGATGCATTATCT 834
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AATATCCAGAATACGTTATCAACTATTAATGAACAAAGACTGATGCATTATCT 480
```

```
CC   DNA to produce fusion proteins with an N-terminal bacterial enzyme
CC   sequence. See also AAQ04686 and AAQ5649-53. (Updated on 10-MAR-2003 to
CC   add missing OS field.)
XX
SQ   Sequence 853 BP; 329 A; 138 C; 128 G; 258 T; 0 U; 0 Other;

Query Match             63.1%;  Score 824.2;  DB 1;  Length 853;
Best Local Similarity   97.9%;  Pred. No. 1.2e-124;
Matches  835;  Conservative  0;  Mismatches 18;  Indels 0;  Gaps 0;
```

```
Qy   163 TTAGTCATCTTTAAGATAATAAAATATCTTAATATTCATGAATAAGAAAAGAATCATC  222
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTAGTCATCTTTAAGATAATAAAATATCTTAATATTCATGAATAAGAAAAGAATCATC   60

Qy   223 TTAAAGACTATTAGTTTGTTAGTGTACAACATCCTTCTTAGCATTGGGATTTCAGCTGT  282
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTAAAGACTATTAGTTTGTTAGTGTACAACATCCTTCTTAGCATTGGGATTTCAGCTGT  120

Qy   283 ATGTCTATTACTAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAGCAGCG   342
         ||||||||||||||||||||||| ||||||||||||||||||||||||| ||||||||
Db   121 ATGTCTATTACTAAAAAGACGCAAACCCAATAATGGCCAAACCCAATTACAAGCAGCG   180

Qy   343 CGAATGAGTAACAGATCAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGAC   402
         ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   181 CGAATGAGTAACTGATCAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGAC   240

Qy   403 TATGCCAAGATGAAGTGAGTAGTTATCATCGTGCTTATGTGGAAGCTGAAACAGTTAACAAT  462
         |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TATGCTAAGATGAAGTGAGTAGTTATCATCGTGCTTATGTGGAAGCTGAAACAGTTAACAAT  300

Qy   463 AACCTTAATGCAACATTAGAACAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATC  522
         |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   301 AACCTTAATGCAACACTAGAACAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATC  360

Qy   523 AACCAAGCTAATACGGGATAAAACGACTTTGATAATGAATAATGAACACCCAAATTAGTTGAAGCA  582
         ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   361 AACCAAGCTAATACGGGATAAAACGACTTTGATAATGAATAATGAACATCCAAATTAGTTGAAGCA  420

Qy   583 TACAAAGCACTAAAAACCACTTTAGAACAACAACGTGCTACTAACCTTGAAGGTTGTCATCA  642
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   421 TACAAAGCACTAAAAACCACTTTAGAACAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCA  480

Qy   643 ACTGCTTATATCAAATTCGGAATAATCAAATTAGTGGATCTATCAATAAAGCTAGTAGTTTA  702
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACTGCTTATATCAAATTCGGAATAATCAAATTAGTGGATCTATCAATAAAGCTAGTAGTTTA  540

Qy   703 ATAACTAAAACACTAGATCCACTAAATGGGGAAGCGCTTTTAGATTCCTAATGAGATTACT  762
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATAACTAAAACACTAGATCCACTAAATGGGGAAGCGCTTTTAGATTCTAATGAGATTACT  600

Qy   763 ACAGTTAATCGGAATTAATAATTAATAACGTTATCAACTATTAATGAACAAAAGACTAATGCT  822
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACAGTTAATCGGAATTAATAATTAATAACGTTATCAACTATTAATGAACAAAAGACTAATGCT  660

Qy   823 GATGCATTATCATCTAATAGTTTTATTAGTAGTTTTATTAAAAAGTGATTCAAAGTTTTGTA  882
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GATGCATTATCATCTAATAGTTTTATTAGTAGTTTTATTAAAAAGTGATTCAAAGTTTTGTA  720

Qy   883 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCT  942
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCT  780

Qy   943 GATGTAACACCCGTCAATTATAAAATGCAAGAAGGACCCGTTTGAAATGGTTGAAGAACCT 1002
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATGTAACACCCGTCAATTATAAAATGCAAGAAGGACCCGTTTGAAATGGTTGAAGAACCT  840

Qy  1003 TCAAGTAGAATTC 1015
         |||||||||||||
Db   841 TCAAGTAGAATTC  853
```

```
Qy   835 AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA  894
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA  540

Qy   895 AACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCTGATGTAACACCC  954
         ||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db   541 AACGCTAATGTTCAACCTAGCCAGTACAGTTTTGTTGCTTTTAGTGCTGATGTAACACCC  600

Qy   955 GTCAATTATAAAATGCAAGAAGGACCCGTTTGAATGTGATGAACCTTCAAGTAGAATT 1014
         ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   601 GTCAATTATAAAATGCAAGAAGGACCCGTTTGAATGTGATGAACCTTCAAGTAGAATT  660

Qy  1015 CTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAAC 1074
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAAC  720

Qy  1075 ACGAAGTACCACAATTAGTTTTAGCAACATGGTCCATCAACTGGTTATTTATATTTCCCT 1134
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACGAAGTACCACAATTAGTTTTAGCAACATGGTCCATCAACTGGTTATTTATATTTCCCT  780

Qy  1135 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAAT 1194
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAAT  840

Qy  1195 GGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATTACACAGTTAAT 1254
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATTACACAGTTAAT  900

Qy  1255 CCAACTTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CCAACTTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT  952
```

```
RESULT 11
AAQ04687
ID   AAQ04687 standard; DNA; 853 BP.
XX
AC   AAQ04687;
XX
DT   10-MAR-2003  (revised)
DT   08-OCT-1990  (first entry)
XX
DE   DNA encoding TMG-1 antigen.
XX
KW   Mycoplasma gallisepticum; poultry; vaccine; ss.
XX
OS   Mycoplasma gallisepticum.
XX
PN   JP02111795-A.
XX
PD   24-APR-1990.
XX
PP   02-JUN-1989;  89JP-00140283.
XX
PR   02-JUN-1988;  88JP-00136343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Kodama K, Saito S, Yanagida N, Kanogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAR05082.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
XX
PS   Claim 5; Fig 2; 20pp; Japanese.
XX
CC   The DNA can be inserted into an expression vector for the prodn. of
CC   polypeptide which elicits an antigen-antibody reaction with anti-
CC   mycoplasma gallisepticum poultry sera. It may also be ligated to other
```

```
RESULT 12
AAN92574
```

```
ID   AAN92574 standard; DNA; 853 BP.
XX
AC   AAN92574;
XX
DT   25-MAR-2003 (revised)
DT   11-MAY-1990 (first entry)
XX
DE   DNA sequence of TM-1 encoding polypeptide TMG-1.
XX
KW   Mycoplasma gallisepticum; Poultry vaccine; ss.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key            Location/Qualifiers
FT   CDS            40..822
FT                  /*tag= a
XX
PN   EP345021-A.
XX
PD   06-DEC-1989.
XX
PF   31-MAY-1989;   89EP-00305441.
XX
PR   02-JUN-1988;   88JP-00136343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI SEIYAKU KK.
XX
PI   Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAP93959.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
XX
PS   Disclosure; Fig 2; 31pp; English.
XX
CC   This base sequence of TM-1 encodes the TMG-1 polypeptide which has the
CC   same sequence as that of a polypeptide expressed in Mycoplasma
CC   gallisepticum in nature. When inserted into a recombinant vector used to
CC   transform a host the antigen protein produced can be used as a vaccine to
CC   prevent and diagnose MG infection. (Updated on 25-MAR-2003 to correct PF
CC   field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
CC   2003 to correct PI field.)
XX
SQ   Sequence 853 BP; 330 A; 139 C; 128 G; 256 T; 0 U; 0 Other;
```

Query Match　　　　61.4%;　　Score 801.8; DB 1; Length 853;
Best Local Similarity　 96.2%;　　Pred. No. 5e-121;
Matches 821;　 Conservative　 0;　Mismatches　 32;　Indels　 0;　Gaps　 0;

```
Qy    163 TTAGTCATCTTTTAAGATATAAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATC 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTAGTCATCTTTTAAGATATAAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATC 60

Qy    223 TTAAAGACTATTAGTTTGTTAGGTACAACATCCCTTCTTAGCATTGGGATTTCTAGCTGT 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTAAAGACTATTAGTTTGTTAGGTACAACATCCCTTCTTAGCATTGGGATTTCTAGCTGT 120

Qy    283 ATGTCTATTACTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCG 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATGTCTATTACTACTAAAAAAGACGCAAACCCAAATAATGGCCAAACCCAATTACAGCAGCG 180

Qy    343 CGAATGAGTAACAGATCTAATCAATGCTAAAGCCATGACATTAGCTTCACTACAAGAC 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGAATGAGTAACTGATCTAATCAATGCTAAAGCCATGACATTAGCTTCACTACAAGAC 240

Qy    403 TATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TATGCTAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 300

Qy    463 AACCTTAATGCAACATTAGAACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATC 522
```

```
Db    301 AACCTTAATGCAACACTAGAACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATC 360
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Qy    523 AACCAAGCTAATACGGATAAAACGACTTTGATAATGAACAACCAAATTAGTTGAAGCA 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AACCAAGCTAATACGGATAAAACGACTTTGATAATGAACAACCAAATTAGTTGAAGCA 420

Qy    583 TACAAAGACTAAAAACCACTTAGAACAAGTGCTACTAACCTTGAAGGTTTGTCATCA 642
              |||||||||||||||||||||||||||||||||||| ||||||||||||||| ||
Db    421 TACAAAGACTAAAAACCACTTAGAACAAGTGCTACTAACCTTGAAGGTTTAGCTTCA 480

Qy    643 ACTGCTTATAATCAAATTCGCAATAATTAGTTGGATCTATACAATACAATAAAGCTAGTAGTTTA 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACTGCTTATAATCAGATTCGTAATAATTAGTTGGATCTATACAATAATGCTAGTAGTTTA 540

Qy    703 ATAACTAAAACACTAGATCCACTAAAATGGGGAACGCTTTTAGATTCTAATGAGATTACT 762
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATAACTAAAACACTAGATCCACTAAAATGGGGAATGCTTTTAGATTCTAATGAGATTACT 600

Qy    763 ACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAATAATGAACTAATGCT 822
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACAGTTAATCGGAATAATAATACGTTATCAACTATTAATGAACAACAAAAGACTAATGCT 660

Qy    823 GATGCATTATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAGTTTTGTA 882
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    661 GATGCATTATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAGTTTTGTA 720

Qy    883 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCT 942
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCT 780

Qy    943 GATGTAACACCCGTCAATTATAAATATGCAAGAAGAACCGTTTGGAATGGTGATGAACCT 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATGTAACACCCGTCAATTATAAATATGCAAGAAGAACGGTTTGAAATGGTGATGAACCT 840

Qy   1003 TCAAGTAGAATTC 1015
              ||||||||||||
Db    841 TCAAGTAGAATTC 853
```

```
RESULT 13
AAN92568
ID   AAN92568 standard; DNA; 708 BP.
XX
AC   AAN92568;
XX
DT   25-MAR-2003 (revised)
DT   11-MAY-1990 (first entry)
XX
DE   DNA sequence of M1 encoding Mycoplasma gallisepticum (MG-1) polypeptide.
XX
KW   Mycoplasma gallisepticum; Poultry vaccine; ss.
OS   Mycoplasma gallisepticum.
XX
PN   EP345021-A.
XX
PD   06-DEC-1989.
XX
PF   31-MAY-1989;   89EP-00305441.
PR   02-JUN-1988;   88JP-00136343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI SEIYAKU KK.
XX
PI   Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAP93646.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
```

XX
PS   Disclosure; Fig 1a; 31pp; English.
XX
CC   This base sequence of M1 encodes the MG1 polypeptide which elicits an
CC   antigen-antibody reaction with anti-MG poultry sera. When inserted into a
CC   recombinant vector used to transform a host the antigen protein produced
CC   can be used as a vaccine to prevent and diagnose MG infection. (Updated
CC   on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct
CC   PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 708 BP; 278 A; 119 C; 107 G; 204 T; 0 U; 0 Other;

        Query Match   52.1%;   Score 679.8; DB 1; Length 708;
        Best Local Similarity 97.6%;   Pred. No. 2.7e-101;
        Matches 690; Conservative   0; Mismatches 17; Indels   0; Gaps   0;

Qy   280 TGTATGTCTATTACTAAAAAGATGCAAACCCAATTAATGGCCAAACCCAATTAGAAGCA 339
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGTATGTCTATTACTAAAAAGACGCAAACCCAATTAATGGCCAAACCCAATTACAAGCA 60

Qy   340 GCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAA 399
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGGCGAATGGAGTTAACTCAATCAATGCTAAAGCAAGACATTAGCTTCACTACAA 120

Qy   400 GACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATATAGTGAAGCTGAAACAGTTAAC 459
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACTATGCTAAGATTGAAGCTAGTTATCATCTGCTTATATAGTGAAGCTGAAACAGTTAAC 180

Qy   460 AATAACCTTAATGCAACATTAGAACACAATTAGAAATGGCTAAAAATGGCTAAAACTAATTTGAATCAGCC 519
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATAACCTTAATGCAACCACTAGAACAACAACTAATTTGAAACTAATTTGAATCAGCC 240

Qy   520 ATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAA 579
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAA 300

Qy   580 GCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGGAGGTTTGTCA 639
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGGAGGTTTAGCT 360

Qy   640 TCAACTGCTTATTAATCAAATTCGCAATAATTAGTGGATCATATACAATAAAAGCTAGTAGT 699
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAACTGCTTATTAATCAGATTCGTAATAATTAGTGGATCATATACAATAAAATGCTAGTAGT 420

Qy   700 TTAATAACTAAAAACACTAGATCCACTAAAATAATACGTTTTTAGATTCTTAATGAGATT 759
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTAATAACTAAAACACTAGATCCACTAAAATGGGGGAACGCTTTTAGATTCTTAATGAGATT 480

RESULT 14
AAQ04686
ID   AAQ04686 standard; DNA; 708 BP.
XX
AC   AAQ04686;
XX
DT   10-MAR-2003 (revised)
DT   08-OCT-1990 (first entry)

XX
DE   DNA encoding MG-1 antigen.
XX
KW   Mycoplasma gallisepticum; poultry; vaccine; ss.
OS   Mycoplasma gallisepticum.
XX
PN   JP02111795-A.
XX
PD   24-APR-1990.
XX
PF   02-JUN-1989;   89JP-00140283.
XX
PR   02-JUN-1988;   88JP-00135343.
XX
PA   (JAPG) NIPPON ZEON KK.
PA   (SHIO) SHIONOGI & CO LTD.
XX
PI   Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAR05081.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
XX
PS   Claim 3; Fig 1a; 20pp; Japanese.
XX
CC   The DNA can be inserted into an expression vector for the prodn. of
CC   polypeptide which elicits an antigen-antibody reaction with anti-
CC   mycoplasma gallisepticum poultry sera. It may also be ligated to other
CC   DNA to produce fusion proteins with an N-terminal bacterial enzyme
CC   sequence. See also AAQ04687 and AAQ05649-53. (Updated on 10-MAR-2003 to
CC   add missing OS field.)
XX
SQ   Sequence 708 BP; 279 A; 119 C; 107 G; 203 T; 0 U; 0 Other;

        Query Match   51.9%;   Score 678.2; DB 1; Length 708;
        Best Local Similarity 97.5%;   Pred. No. 5e-101;
        Matches 689; Conservative   0; Mismatches 18; Indels   0; Gaps   0;

Qy   280 TGTATGTCTATTACTAAAAAGATGCAAACCCAATTAATGGCCAAACCCAATTAGAAGCA 339
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGTATGTCTATTACTAAAAAGACGCAAACCCAATTAATGGCCAAACCCAATTACAAGCA 60

Qy   340 GCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAA 399
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGGCGAATGGAGTTAACTCAATCAATGCTAAAGCAAGACATTAGCTTCACTACAA 120

Qy   400 GACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATATAGTGAAGCTGAAACAGTTAAC 459
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACTATGCTAAGATTGAAGCTAGTTATCATCTGCTTATATAGTGAAGCTGAAACAGTTAAC 180

Qy   460 AATAACCTTAATGCAACATTAGAACACAATAAAATGGCTAAAAATGGCTAAAAACTAATTTAGAATCAGCC 519
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATAACCTTAATGCAACCACTAGAACAACAACTAAAAATGGCTAAAAATAATGGCTAAAAACTAATTTAGAATCAGCC 240

Qy   520 ATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAA 579
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACATCCCAAATTTAGTTGAA 300

Qy   580 GCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCA 639
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCT 360

Qy   640 TCAACTGCTTATATCAAATTCGCAATAAATTTAGTGGATCATATACAATAAAAGCTAGTAGT 699
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAACTGCTTATATCAATCAGATTCGTAATAATTAGTGGATCTATACAATAATAATGCTAGTAGT 420

Qy   700 TTAATAACTAAAACACTAGATCCACTAAAATAATACGTTTTTAGATTCTTAATGAGATT 759
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTAATAACAAAAACACTAGATCCACTAAAATGGGGGAATGCTTTTAGATTCTTTAGATTCAATGAGATT 480

```
CC   sequence, and a probe based on nucleotides 718-41 of AAT04076. The method
CC   using these sequences is faster i.e. is able to detect M.gallisepticum,
CC   which causes avian chronic respiratory diseases, within one day, from
CC   avian secretions, washings from the lung, respiratory tract, nasal
CC   cavity, etc
XX
SQ   Sequence 2196 BP; 800 A; 353 C; 376 G; 667 T; 0 U; 0 Other;

Query Match             16.6%;  Score 216.6;  DB 2;  Length 2196;
Best Local Similarity   55.1%;  Pred. No. 2.3e-26;
Matches  548;  Conservative  0;  Mismatches 414;  Indels  33;  Gaps   5;

Qy   329  AATTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAG  388
          ||||| ||  || ||  |||| || |  ||||| | ||  || |||| ||  ||||| |||
Db   352  AATTAGCAGACGCAAAAGCTGCTTTAACTACTTGATTATGGTGAAACTGCAAATCTTG   411

Qy   389  CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTG   448
          | ||| | | ||||||||| |||||  |||| |||  | ||| | || |||| | |||||
Db   412  CGTCATATGAAGACTATGCTAAGATCAAAAGTGAATTAACATCAGCGTATGACAGCTA   471

Qy   449  AAACAGTAACAATAACCTTAATGCAACAGTTAGAGACAACTAAAAATGGCTAAAACTAATT  508
          || ||||  | |  || | | || || |  ||||| ||||| || |  |||||||||||| 
Db   472  AAGCAGTTCAGTAAAACTCGGTGCAACTCGTAAATCGAGGTTAATGAGGCAAAAACTACAT  531

Qy   509  TAGAATCAGCCATCAACCAAGCTAATACGGATAAAAACAGGACTTTGATAATGAACACCCAA  568
          ||||| | || || |  |||||  ||  |||  || |||| | |||| |||||||| |||||
Db   532  TAGATGCTGCTATTAAAAAGCTGCCAGTGCTAAGAACAACCACTTAGAACACAACGTGCACAGCACGGGT  591

Qy   569  ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTAGAACAACGTGCACAGCACGGGT  628
          | | || ||||||||| || | ||| |||| ||| ||  || | |   | ||   ||||
Db   592  CACTAGTGGAAGCATATAACAATCGTAAAAGAACGTAAAAGAAGAAAAACTAATTTAG   651

Qy   629  AAGGTTTGTCATCAACTGCTTATAATCAATCAAATTCGCAATAATTAGTGATCTATACAATA  688
          |  | || | |  || |||| |  ||| | |||||   ||  | || | || |||| |||||
Db   652  ATTCTCTGCAAAGCAAAATATGCAGCAATCAGAACTAATCTTATAGTTTTATATGAAA    711

Qy   689  AAGCTAGTAGTTTAATAACTAAAAACTACTGATCCACTAAATGGGGAACGCTTTAGATT   748
          |||| | | || ||| | | || ||||||  |   |||  | ||||| | |||| ||| 
Db   712  AAGCCAATACATTGTTACAGTCGTACTTTAGAGCCCTGCTACTGGAAATATTCCTGAAGTTA  771

Qy   749  CTAATGAGATTACTACAGTTAATCGGAATAATTAAATACGTTATCAACTATTAATGAAC   808
          ||||||  || |||||| |  | |    |  | ||| | |||| |||  | |||||||| 
Db   772  TGAGTG---TAACACAAGCTAATCAAGATATTACTACTCAAGACTTCAAGATTAATAATGGCTT  828

Qy   809  AAAAGACTAATGCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATG   868
          ||||  | ||||| || | |||| ||| |||||||| | |  | | ||| ||| |||| 
Db   829  GAAAACAAAATGCTGATAAATAATCTGATAATTTTATCTAACAGTCTTTAGTTAAAAATA   888

Qy   869  AACAAAGTTTTGTAGGGACTTTACAAACGTTCTTTACAACGTTCAACCTTCAAACTACAGTTTG  928
          |   | ||| |  | |  || ||  || || |  |  | ||| ||   |||   |||||||||| 
Db   889  ATTTGACTAGAGTTG---ATGTAGCAAATAATCAGGAGCAACCAGCAAATTACAGTTTTG   945

Qy   929  TTGCTTTTAGTGCTGATGTAACACCCGTCAATGTAACTAACTATAATAAATGCAAGAAGGACCGTTGGA  988
          |||  |||||||  | |||  |   ||  || ||| ||||||| | ||||| |  ||||| | ||| ||
Db   946  TTGGTTTTAGTGTTAATGTTGATACCCTAACTAACTAGTAGTTTAAATGTGCGCAAAGAAAGTTTGGG  1005

Qy   989  ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTA-----------  1033
          | | |||||||| || || |||   |  |||||| | || | |  
Db  1006  CCTCTGAAAATACTCCTTTAGCAACTACACCAGCTGAAGATGCAACCACAAGAAGTGCAT  1065

Qy  1034  ---TCAACGATGTTTCTGATGTAACCGTCAATTATAGTTATAGTTCTGAACGTGTGCGTTA  1090
             |  ||||||||||| ||| || | || |||| | |||| | |||| || ||| | |||
Db  1066  CCTAACGATGTTTCATGAATCATATAGTAGTAGTTTAAATGTGTGCAAAGAAAGTAAATCA  1125

Qy  1091  GTTTTAGCAACTATGG---TCCATCAACTGGTTATTATATTCCCTTATAAGTTGGTTA    1147
          | |||  || || |||   |||||| || | |||  | ||||||||| || ||| |||
Db  1126  GCTTTCGTTACTTTGGAGCTGAAAAAACAGCTTACTTTATATTTCCCTATAAATTAGTTA  1185

Qy  1148  AAGCAGCTGATGCTAAACCGTTGGATTACAATAACAATAACAAATAAATGGAAATGTTCAAC  1207
          ||                   |  ||| |||||  ||| | | |||||||| ||  | ||||
Db  1186  AAAC------TAGTGATAATGTGGTTTACAATATAGTAAGTAAAATGGTGGTGATACTAAAC  1239

Qy  1208  AAGTTGAGTTTGCCACTTCACTAGTGCAAATAAATACTACAGCTAATCCAACTCCAGCAG  1267
```

```
Qy   760  ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAAT   819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAAT   540

Qy   820  GCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTT    879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   541  GCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTT    600

Qy   880  GTAGGGACTTTACAAACGCTAATGTTCAAACCTTCAAACTACAGTTTGTTGCTTTTTAGT   939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GTAGGGACTTTACAAACGCTAATGTTCAAACCTTCAAACTACAGTTTGTTGCTTTTTAGT   660

Qy   940  GCTGATGTAACACCCGTCAATTATAAAATATGCAAGAAGGACCGTTTG  986
          ||||||||||||||||||||||||||||| || |||||||  ||||||
Db   661  GCTGATGTAACACCCGTCAATTATAAAATATGCAAGAAGGAACGTTTG  707
```

```
RESULT 15
AAT04075
ID   AAT04075 standard; DNA; 2196 BP.
XX
AC   AAT04075;
XX
DT   19-JUL-1996  (first entry)
XX
DE   M.gallisepticum DNA sequence I encodes 661 amino acid protein.
XX
KW   Detection; probe; primer; PCR; amplification; secretion; lung;
KW   avian chronic respiratory disease; respiratory tract; nasal cavity; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             171..2156
FT                   /*tag= a
FT   misc_difference 828..830
FT                   /*tag= b
FT                   /codon= seq: TGA, aa: Trp
FT   misc_difference 978..980
FT                   /*tag= c
FT                   /codon= seq: TGA, aa: Trp
FT   misc_difference 1083..1085
FT                   /*tag= d
FT                   /codon= seq: TGA, aa: Trp
FT   misc_difference 1911..1913
FT                   /*tag= e
FT                   /codon= seq: TGA, aa: Trp
XX
PN   JP0723649B-A.
XX
PD   12-SEP-1995.
XX
PF   25-FEB-1994;  94JP-00052764.
XX
PR   25-FEB-1994;  94JP-00052764.
XX
PA   (JAPG )  NIPPON ZEON KK.
PA   (SHIO )  SHIONOGI & CO LTD.
XX
DR   WPI; 1995-347462/45.
DR   P-PSDB; AAR79910.
XX
PT   Detection of Mycoplasma gallisepticum - for the quick detection, i.e.
PT   within one day, of avian chronic respiratory diseases.
XX
PS   Claim 2; Page 7-10; 11pp; Japanese.
XX
CC   This is the nucleotide sequence of a fragment of the Mycoplasma
CC   gallisepticum genome which codes for a 661 amino acid protein. This
CC   sequence and the sequence of AAT04076 (encoding a 261 amino acid protein)
CC   can be used to detect M.gallisepticum using probes based on nucleotides
CC   1125-1648 and primers based on nucleotides 449-466, the complement of
CC   bases 893-919, 1908-1934 and the complement of bases 2184-2210 of this
```

```
Db        1240 AAATTAACTTTGTACAAACTCCAGCTTCTCTGGTTCAAGTGATGATGTTGCTGCTAATGAAGAAG 1299
                    ||  ||| | |||||              |     |      |   | |    ||  ||
Qy        1268 TTGATGAGATTAAAGTTGCTAAAATCGTTTTATCA 1302
                    |     |     |||  |||  ||| ||
Db        1300 AAACTATGGCTAGTCCTGCTGAAATGCAGTCAGCA 1334
```

Search completed: May 5, 2004, 12:25:44
Job time : 463.318 secs

OM nucleic - nucleic search, using sw model

```
            GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:     May  5, 2004, 10:52:31 ; Search time 98.4948 Seconds
                                      (without alignments)
                                      7358.420 Million cell updates/sec

Title:      US-09-901-572A-1
Perfect score:  1306
Sequence:   1 aaaaacatcagattgttaat........taaaatcgtttatcaggtt 1306

Scoring table:  IDENTITY.NUC
            Gapop 10.0 , Gapext 1.0

Searched:   682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents_NA:*
            1: /cgn2_6/ptcdata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptcdata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptcdata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptcdata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptcdata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptcdata/2/ina/backfile1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1280.4	98.0	2144	2	US-08-525-742-9	Sequence 9, Appli
	2	1270	97.2	1387	2	US-08-525-742-3	Sequence 3, Appli
	3	1264	96.8	1305	1	US-08-185-851A-1	Sequence 1, Appli
	4	1086.8	83.2	2014	2	US-08-525-742-7	Sequence 7, Appli
	5	216.6	16.6	2359	2	US-08-525-742-1	Sequence 1, Appli
c	6	66.4	5.1	6669	4	US-10-204-708-6	Sequence 6, Appli
c	7	65.6	5.0	3057	4	US-09-601-198-55	Sequence 55, Appl
c	8	61.2	4.7	5152	4	US-10-204-708-73	Sequence 73, Appl
c	9	59.4	4.5	11049	4	US-10-204-708-23	Sequence 23, Appl
c	10	55	4.2	10640	4	US-09-417-485D-5	Sequence 5, Appli
c	11	54	4.1	640681	4	US-09-790-988-1	Sequence 1, Appli
c	12	53.8	4.1	5476	4	US-10-204-708-82	Sequence 82, Appl
c	13	53.6	4.1	1887	4	US-09-601-198-39	Sequence 39, Appl
c	14	53.2	4.1	6113	4	US-10-204-708-14	Sequence 14, Appl
c	15	53.2	4.1	6317	4	US-10-204-708-12	Sequence 12, Appl
c	16	53.2	4.1	8700	2	US-08-645-193B-18	Sequence 18, Appl
c	17	52.8	4.0	7304	4	US-10-204-708-44	Sequence 44, Appl
c	18	52.4	4.0	429	1	US-09-543-681A-707	Sequence 707, App
c	19	52.4	4.0	1575	4	US-09-543-681A-615	Sequence 615, App
c	20	51.6	4.0	8700	2	US-08-392-625-16	Sequence 16, Appl
c	21	51.6	4.0	8700	2	US-08-466-961A-16	Sequence 16, Appl
c	22	51.6	4.0	53332	4	US-09-801-861-3	Sequence 3, Appli
c	23	51.2	3.9	6156	4	US-10-204-708-59	Sequence 59, Appl
c	24	51.2	3.9	6861	4	US-10-204-708-20	Sequence 20, Appl
c	25	50.8	3.9	1851	4	US-09-601-198-51	Sequence 51, Appl
c	26	50.6	3.9	14066	4	US-09-601-198-56	Sequence 56, Appl
c	27	50.2	3.8	7664	4	US-10-204-708-83	Sequence 83, Appl
c	28	50	3.8	2043	4	US-09-601-198-181	Sequence 181, App
c	29	49.8	3.8	8093	4	US-10-204-708-31	Sequence 31, Appl
c	30	49.4	3.8	615	3	US-08-998-416-186	Sequence 186, App
c	31	49.4	3.8	6866	4	US-10-204-708-20	Sequence 20, Appli
c	32	49.4	3.8	11049	4	US-10-204-708-21	Sequence 21, Appli
c	33	49.2	3.8	837	2	US-08-998-416-288	Sequence 288, App
	34	49.2	3.8	8920	3	US-08-446-855A-1	Sequence 1, Appli
	35	49.2	3.8	8920	3	US-09-150-741-1	Sequence 1, Appli
	36	49	3.8	640681	4	US-09-790-988-1	Sequence 80, Appl
c	37	49	3.8	8961	4	US-10-204-708-80	Sequence 80, Appl
c	38	48.6	3.7	636	3	US-08-998-416-1137	Sequence 1137, Ap
c	39	48.6	3.7	8607	4	US-10-204-708-72	Sequence 72, Appl
c	40	48.2	3.7	7304	4	US-10-204-708-43	Sequence 43, Appl
c	41	48.2	3.7	1500	4	US-09-601-198-36	Sequence 36, Appl
	42	48.2	3.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	43	48	3.7	6317	2	US-10-204-708-11	Sequence 11, Appl
c	44	48	3.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	45	47.8	3.7	6070	4	US-10-204-708-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-08-525-742-9
; Sequence 9, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
;   APPLICANT: Saito, Shuji
;   APPLICANT: Ohkawa, Setsuko
;   APPLICANT: Saeki, Sakiko
;   APPLICANT: Ohsawa, Ikuroh
;   APPLICANT: Funato, Hirono
;   APPLICANT: Iritani, Yoshikazu
;   APPLICANT: Aoyama, Shigemi
;   APPLICANT: Takahashi, Kiyoohito
;   TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
;   TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;   TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;   TITLE OF INVENTION: AS USE THEREOF
;   NUMBER OF SEQUENCES: 51
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;     ADDRESSEE: NAUGHTON
;     STREET: 1725 K Street, Suite 1000
;     CITY: Washington
;     STATE: DC
;     COUNTRY: USA
;     ZIP: 20006
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/525,742
;     FILING DATE: 25-SEP-1995
;     CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 05-074139
;     FILING DATE: 31-MAR-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 05-245625
;     FILING DATE: 30-SEP-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/JP94/00541
;     FILING DATE: 31-MAR-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: McLeland, Le-Nhung
;     REGISTRATION NUMBER: 31,541
;     REFERENCE/DOCKET NUMBER: 950811
;     TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-659-2930
```

; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 202..2046
US-08-525-742-9

Query Match 98.0%; Score 1280.4; DB 2; Length 2144;
Best Local Similarity 98.8%; Pred. No. 2.2e-262;
Matches 1290; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 AAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAAATCTTCTAACAA 60
Db 1 AAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAAATCTTCTAACAA 60

Qy 61 AATCCTAAATAAATAAGCCGTTAAATAACTAACTAAAAAAAATTAAAAAAATGGTTTTTCTTATC 120
Db 61 AATCCTAAATAAATAAGCCGTTAAATAACTAACTAAAAAAAATTAAAAAAATGGTTTTTCTTATC 120

Qy 121 AACCAAAATTCTCCAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGAT 180
Db 121 AACCAAAATTCTCCAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGAT 180

Qy 181 ATAAATATATCTTAATATTCTATGAATAAGAAAAAGAATCATCTTAAAGACTATTAGTTTG 240
Db 181 ATAAATATATCTTAATATTCTATGAATAAGAAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy 241 TTAGGTACAACATCCTTTCTTGCTAGCATTGGGATTTCTAGTGTATGTCTATTACTAAAAA 300
Db 241 TTAGGTACAACATCCTTTCTTGCTAGCATTGGGATTTCTAGTGTATGTCTATTACTAAAAA 300

Qy 301 GATGCAAACCCAAATAAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGAT 360
Db 301 GACGCAAACCCAAATAAATGGCCAAACCCAATTACAAGCAGCGCGAATGGAGTTAACTGAT 360

Qy 361 CTAATAGCTAAAGCGATGACATTAGCTCCACTACAAGACTATGCCAAGATTGAAGCT 420
Db 361 CTAATCAATGCTAAAGCCAAGGACAATTAGCTTCCACTACAAGACTATGCCATGCTAAGAT 420

Qy 421 AGTTTATCATCTGCTTATATGTGAAGCTGAAACAGTTAACAATAACCTTAAATGCAACATTA 480
Db 421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAAATGCAACACTA 480

Qy 481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAGCTAATACGGAT 540
Db 481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAGCTAATACGGAT 540

Qy 541 AAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTCAAAACC 600
Db 541 AAAACGACTTTTGATAATGAACATCCAAATTAGTTGAAGCATACAAAGCACTCAAAAACC 600

Qy 601 ACTTTAGAACAACGTGCTACTACAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
Db 601 ACTTTAGAACAACGTGCTACTACAACCTTGAAGGTTTAGCTTCAACTGCTTATAATCAGATT 660

Qy 661 CGCAATAATTAGTTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAAACACTAGAT 720
Db 661 CGTAATAATTAGTTGGATCTATACAATAATGCTAGTAGTTTAATAACTAAAAACACTAGAT 720

Qy 721 CCACTAAATGGGGGAACGCTTTAGATTCTAAATGAGATTACTACAGTTAATCGGAATATT 780
Db 721 CCACTAAATGGGGGAATGCTTTAGATTCTAAATGAGATTACTACAGTTAATCGGAATATT 780

Qy 781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTAATGCGATTATCTAATAGT 840
Db 781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCGATTATCTAATAGT 840

Qy 841 TTTATTAAAAAGTGATTCAAATAATGAACAAGTTTGTAGGGACTTTTACAAACGCT 900
Db 841 TTTATTAAAAAGTGATTCAAATAATGAACAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy 901 AATGTTCAACCTCAAACTACACAGTTTTGTCTTTTAGTGCTGATGTAACACCCGTCAAT 960
Db 901 AATGTTCAACCTCAAACTACACAGTTTTGTCTTTTAGTGCTGATGTAACACCCGTCAAT 960

Qy 961 TATAAATATGCAAGAAGAGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
Db 961 TATAAATATGCAAGAAGAGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020

Qy 1021 AACACGAATAGTATCACAGATGTTCTTGGATTATAGTTAGCTGGAACAAACACGAAG 1080
Db 1021 AACACGAATAGTATCACAGATGTTCTTGGATTATAGTTAGCTGGAACGAACACGAAG 1080

Qy 1081 TACCAATTAGTTTAGCAACTAGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
Db 1081 TACCAATTAGTTTAGCAACTAGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140

Qy 1141 TTGGTTAAAGCAGCTGAATCGTAATAACGTTGGATTACAATACAAATTAAATAATGAAAT 1200
Db 1141 TTGGTTAAAGCAGCTGAATCGTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy 1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATATATACTACAGCTAATCCAACT 1260
Db 1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATATATACTACAGCTAATCCAACT 1260

Qy 1261 CCAGCAGTTGATGAGTTAAGTTGCTAAAATCGTTTTATCAGGTT 1306
Db 1261 CCAGCAGTTGATGAGTTAAAAGTTGCTAAAATCGTTTTATCAGGTT 1306

RESULT 2
US-08-525-742-3
; Sequence 3, Application US/0852574 2
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyoohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625

```
; FILING DATE:      30-SEP-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  PCT/JP94/00541
;   FILING DATE:      31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
;   NAME:  McLeland, Le-Nhung
;   REGISTRATION NUMBER:  31,541
;   REFERENCE/DOCKET NUMBER:  950811
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  202-659-2930
;   TELEFAX:    202-8870357
; INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1387 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  202..1305
US-08-525-742-3
```

```
Query Match     97.2%;  Score 1270;  DB 2;  Length 1387;
Best Local Similarity   98.8%;  Pred. No. 3.2e-260;
Matches 1290;  Conservative    0;  Mismatches  15;  Indels   1;  Gaps   1;

Qy     1 AAAAACATCAGATTGTTAATCGATATCTTGCTTAAAAAAACACAAAAATCTTCTAACAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AAAAACATCAGATTGTTAATCGATATCTTGCTTAAAAAAACACAAAAATCTTCTAACAA 60

Qy    61 AATCCTAAATAAATAAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AATCCTAAATAAATAAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATC 120

Qy   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTTATTTTATTTTAGTCATCTTTTAAGAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTTATTTTATTTTAGTCATCTTTTAAGAT 180

Qy   181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy   241 TTAGGTACAACATCCTTTCTTAGCATTGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTAGGTACAACATCCTTTCTTAGCATTGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300

Qy   301 GATGCAAACCCAAATAAATGGCCAAAACCCAATTAGAAGCAGCCGAATGGAGTTAACAGAT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   301 GACGCAAACCCAAATAAATGGCCAAAACCCAATTACAAGCAGCCGAATGGAGTTAACTGAT 360

Qy   361 CTAATCAATGCTAAAGGCGGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTAATCAATGCTAAAGGCGGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 420

Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTAAATAATTTAACAATAACCTTAATGCAACATTA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGTTTATCATCTGCTTATAGTGAAGCTAAATAATTTAACAATAACCTTAATGCAACACTA 480

Qy   481 GAACAACTAAAAATGGCTAAAAACTAATTTAGTTGAAGCTAATCAGCTAATACCGGAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAACAACTAAAAATGGCTAAAAACTAATTTAGTTGAAGCTAATCAGCTAATACCGGAT 540

Qy   541 AAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 600

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTAGCTTCAACTGCTTATAATCAGATT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTAGCTTCAACTGCTTATAATCAAATT 660

Qy   661 CGCAATAATTTAGTGGATCTATACAATAATGCTAGTAGTTTAATAACTAAAACACTAGAT 720
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGTAATAATTTAGTGGATCTATACAATAATGCTAGTAGTTTAATAACTAAAACACTAGAT 720

Qy   721 CCACTAAATGGGGGAACGCCTTTAGATTCTTAATGACCAGTAATACTACACAGTTAATCGGAATATT 780
           ||||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db   721 CCACTAAATGGGGGGATGCTTTAGATTC TTAATGAGATTACTACACAGTTAATCGGAATATT 780

Qy   781 AATAACTACGTTATCAACTAATATAAGTAATGAACAAAAGACTAATCTGATGCATTATCTAATAGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AATAATACGTTATCAACTAATATAAGTAATGAACAAAAGACTAATGTGATGCATTATCTAATAGT 840

Qy   841 TTTATTAAAAAGTGATTCAAAATAATGATGAACAAAGTTTGTGAGGGACTTTCAAACGCT 900
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   841 TTTATTAAAAAGTGATTCAAAATAATGATGAACAAAGTTTGTGAGGGACTTTTACAAACGCT 900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTGTTGCTTTTGTTGTCTTTCAAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTGTTGCTTTTGTTGTCAACCCGTCAAT 960

Qy   961 TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCTTCAAGTAGAAATTCTTGCA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCTTCAAGTAGAAATTCTTGCA 1020

Qy  1021 AACACGAATAGTATCACAGATGTTCTTGGATTATAGTTAGCTGGAACAAACACGAAG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AACACGAATAGTATCACAGATGTTCTTGGATTATAGTTAGCTGGAACAAACACGAAG 1080

Qy  1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140

Qy  1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGAAAAT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGAAAAT 1200

Qy  1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATATACTCAGCTAATCCAACT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATATACTCAGCTAATCCAACT 1260

Qy  1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
           |||||||||||||||||||||||||||||||||||||||||||||
Db  1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1305
```

```
;      CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;      NAME: Theresa M. Stevens-Smith
;      REGISTRATION NUMBER: 36,281
;      REFERENCE/DOCKET NUMBER: PO-8-A930918
;  TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 202-659-2930
;      TELEFAX: 202-887-0357
;  INFORMATION FOR SEQ ID NO: 1:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 1305 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: circular
;      HYPOTHETICAL: NO
;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 202..1308
US-08-185-851A-1

Query Match      96.8%;  Score 1264;  DB 1;  Length 1305;
Best Local Similarity  98.3%;  Pred. No. 5.9e-259;
Matches 1284;  Conservative  0;  Mismatches  0;  Indels  21;  Gaps  1;

Qy      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATC 120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTATTTTTATTTTTAGTCATCTTTAAGAT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTATTTTTATTTTTAGTCATCTTTAAGAT 180

Qy    181 ATAAATATATCTTAATATCTTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATAAATATATCTTAATATCTTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy    241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300

Qy    301 GATGCAAACCCAAATAATGCCAAACCCAATTAGAAGCAGAGCGCGAATGGAGTAAACAGAT 360
            || ||||||||||||||||||||||||||| |||||||| ||||||||||||| || |||
Db    301 GACGCAAACCCAAATAATGCCAAACCCAATTACAAGCAGCGCGCGAATGGAGTAACTGAT 360

Qy    361 CTAATCAATGCTAAAGCGATGAAGCTCACTACAAGACTATGCCAAGATTGAAGCT 420
            ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAGCGATGAAGCTCACTACAAGACTATGCCAAGATTGAAGCT 420

Qy    421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGCAGTTAACAATAACCTTAATGCAACATTA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGCAGTTAACAATAACCTTAATGCAACACTA 480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAACTCAGCGCCATCAACCAAGCTAATACGGAT 540
            ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCGCCATCAACCAAGCTAATACGGAT 540

Qy    541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC 600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT 660

Qy    661 CGCAATAATTTAGTGGATCTCATACAGAATAAAGCTAGTAGTTAATAACTAAAAACACTAGAT 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGTAATAATTTAGTGGATCATCATACAGAATAAAGCTAGTAGTTAATAACTAAAAACACTAGAT 720

Qy    721 CCACTAAATGGGGGAATGCTTTAGATTCTAATGATTCTAATGAGATTACTACAGTAATCGGAATATT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCACTAAATGGGGGAATGCTTTAGATTCTAATGATTCTAATGAGATTACTACAGTAATCGGAATATT 780

Qy    781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840

Qy    841 TTTATTAAAAAGTGATTCAAAATAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTATTAAAAAGTGATTCAAAATAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy    901 AATGTTCAACCTTCAAACTACACAGTTTTGTTGCTTTTGTAGTGCTGATGTAACACCCGTCAAT 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AATGTTCAACCTTCAAACTACACAGTTTTGTTGCTTTTGTAGTGCTGATGTAACACCCGTCAAT 960

Qy    961 TATAAAATGCAAGAAGGACCGTTGGAATGGTGATGAAACCTTCAGTAGAATTCTTGCA 1020
            |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db    961 TATAAAATGCAAGAAGGACCGTTNNNAATGGTGATGAAACCTTCAGTAGAATTCTTGCA 1020

Qy   1021 AACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTAGCTGAACAAACACGAAG 1080
            ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db   1021 AACACGAATAGTATCACAGATGTTCTNNNATTATAGTTAGCTGAACAAACACGAAG 1080

Qy   1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140

Qy   1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy   1201 GTTCAACAAGTTGAGTTTGCCACTTGCAACTAGTGCAAATAATATACACAGCTAATCCAACT 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCAACAAGTTGAGTTTGCCACTTGCAACTAGTGCAAATAATATACACAGCTAATCCAACT 1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTTTATCAGGTT 1306
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTTTATCAGGTT 1305
```

RESULT 4
US-08-525-742-7
; Sequence 7, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyoohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742

```
;       FILING DATE: 25-SEP-1995
;       CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 05-074139
;       FILING DATE: 31-MAR-1993
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 05-245625
;       FILING DATE: 30-SEP-1993
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: PCT/JP94/00541
;       FILING DATE: 31-MAR-1994
;   ATTORNEY/AGENT INFORMATION:
;       NAME: McLeland, Le-Nhung
;       REGISTRATION NUMBER: 31,541
;       REFERENCE/DOCKET NUMBER: 950811
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 202-659-2930
;       TELEFAX: 202-8870357
;   INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 2014 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 54..1883
US-08-525-742-7

Query Match          83.2%;     Score 1086.8;   DB 2;   Length 2014;
Best Local Similarity 96.7%;     Pred. No. 2.1e-221;
Matches 1120; Conservative  0; Mismatches  37; Indels   1; Gaps   1;
```

```
Qy     150  TTTATTTTATTTTAGT-CATCTTTTAAGATATAAATATATCTTAATATTCTATGAATA  208
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       1  TTTATTTTATTTTGGTAAATCTTTTAAATATAATATATTTAATATTCTATGAATA  60

Qy     209  AGAAAAGAAATCATCTTAAAGACTATTAGTTTGTTAGTACAACATCCTTCTTAGCATTG  268
            | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  AAAAAGAAATCATCTTAAAGACTATTAGCTTGTTAGTACAACATCCTTCTTAGTATTG  120

Qy     269  GGATTTCTAGCTGTATGTCTATTACTTAAAAAAGATGCAAACCCAAATAATGGCCAAACCC  328
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  GGATTTCTAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCC  180

Qy     329  AATTAGAAGCAGCGCGGAATGGAGTTAACAGATCTAATCAATGCTAAAGCTAAAGCGATGACATTAG  388
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  AATTAGAAGCAGCGCGGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAG  240

Qy     389  CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG  448
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG  300

Qy     449  AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACAACTAAAAATGGCTAAAACTAATT  508
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACAACTAAAAATGGCTAAAACTAATT  360

Qy     509  TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAA  568
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAA  420

Qy     569  ATTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACCGTGCTACTAACCTTG  628
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACCGTGCTACTAACCTTG  480

Qy     629  AAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTAGTGATCTCATACAATA  688
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  AAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTAGTGATCTCATACAATA  540

Qy     689  AAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCTTTTAGATT  748
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  AAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCTTTTAGATT  600
```

```
Qy     749  CTAATGAGATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAAC  808
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  CTAATGAGATTACTACAGTTAATAAGAATATTAATAATACGTTATCAACTATTAATGAAC  660

Qy     809  AAAAGACTAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATG  868
            |||||||||||||||||||||||||||||||||||||||||||||
Db     661  AAAAGACTAATGCTGATGCATTAGCATTAGCTAATAGTTTTATTAAAGAACTGATTCAAAATAATA  720

Qy     869  AACAAAGTTTGTAGGGACTTTACAAACGCTAATGTCCAACCTTCAAACTACAGTTTTG  928
            | |||||||||||||||| |||| |||||||||||||||||||| |||||||||
Db     721  AACAAAGTTTGTAGGAATGTTTACAAACACTAATGTTCAACCTTCAAACTATAGTTTTG  780

Qy     929  TTGCTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGA  988
            ||||||||||||||| ||||||||| |||| | |||||||||||||||||||
Db     781  TTGCTTTAGTGCTGATGTAACACCTGTTAATTAATAAATATGCAAGAAGAACGGTTTGAA  840

Qy     989  ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTT  1048
            |||||||||||||||||||| ||||||||||||||||||||||||||||
Db     841  ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACCAATAGTATTACTGATGTTTCAT  900

Qy    1049  GGATTTATAGTTAGCTGGACAAACACGAGTACCAATTTAGTTTTAGCAACTATGGTC  1108
            | ||||| || |||| ||||||||||||||||| ||||||||||||||||
Db     901  GAATTTATAGTTTATCTGGAACAACAAACACGAAATACCAATTTAGTTTTAGCAACTACGGTC  960

Qy    1109  CATCAACTGGTTATTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACG  1168
            ||||||||||| |||||||||||||||||||||||||||||||||||||
Db     961  CATCAACTGGTTATTTATTTCCCTTATAAGTGGTTAAAGCGGCTGATGCTAGTAATG  1020

Qy    1169  TTGGATTACAATACACAATTAAATAATGGAAAATGTTCAACAGTTGCCACTTCAA  1228
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  TTGGATTACAATACAATAAAACTAAATAATGGAAAATGTTCAACCAGTTGAGTTTGCCACTTCAA  1080

Qy    1229  CTAGTGCAAATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA  1288
            ||||||||||||||||||||||||||||||||||||||||||||
Db    1081  CTAGGCCAAATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA  1140

Qy    1289  AAATCGTTTTATCAGGTT  1306
            |||||||||||
Db    1141  AAATCGTTTTATCAGGTT  1158
```

RESULT 5
US-08-525-742-1
; Sequence 1, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
;   APPLICANT: Saito, Shuji
;   APPLICANT: Ohkawa, Setsuko
;   APPLICANT: Saeki, Sakiko
;   APPLICANT: Ohsawa, Ikuroh
;   APPLICANT: Funato, Hirono
;   APPLICANT: Iritani, Yoshikazu
;   APPLICANT: Aoyama, Shigemi
;   APPLICANT: Takahashi, Kiyoohito
;   TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
;   TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;   TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;   TITLE OF INVENTION: AS USE THEREOF
;   NUMBER OF SEQUENCES: 51
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;   ADDRESSEE: NAUGHTON
;   STREET: 1725 K Street, Suite 1000
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20006
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/525,742
;   FILING DATE: 25-SEP-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 05-074139
;   FILING DATE: 31-MAR-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 05-245625
;   FILING DATE: 30-SEP-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP94/00541
;   FILING DATE: 31-MAR-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: McLeland, Le-Nhung
;   REGISTRATION NUMBER: 31,541
;   REFERENCE/DOCKET NUMBER: 950811
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-659-2930
;   TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2369 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 171..2153
US-08-525-742-1

Query Match   16.6%;   Score 216.6;   DB 2;   Length 2369;
Best Local Similarity   55.1%;   Pred. No. 3.5e-37;
Matches 548;   Conservative   0;   Mismatches 414;   Indels   33;   Gaps   5;

Qy   329 AATTAGAAGCAGCGCGAATGGAGTTAACAGATCCAATCAATGCTAAAGCGATGACATTAG 388
              |||||| || || ||||| || ||| ||| ||| |||||||||
Db   352 AATTAGCACGCAAAAGCTGCTTAACTACTTTGATTAATGGTGAAACTGCAAATCTTG 411

Qy   389 CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG 448
              ||||| || |||| | | ||| ||| ||||| |||| ||| ||||
Db   412 CGTCATATGAAGACTATGCTAGATCAAAAGTGAATTAACATCAGCGTATGAAACAGCTA 471

Qy   449 AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATT 508
              |||||| |||| ||| || |||| |||||| ||| ||| |||||||
Db   472 AAGCAGTTTCAGTCAACTAAAACTGGTGCAACCTTAAATGAGGTTAATGAAGGCAAAAACTACAT 531

Qy   509 TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACAACTTTTGATAATGAACACCCAA 568
              |||| | || | ||| ||||| || |||||||| ||||||| |||||||
Db   532 TAGATGCTGCTATTAAAAAAGCTCGTAGTGCTAAGAATGATTTTGATGCACAGCACGGGT 591

Qy   569 ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAAACGTTGCTACTAACCTG 628
              ||||||| ||| |||| ||| || || || | |||| | ||| | ||||||
Db   592 CACTAGTGAAGCATAACAATCTAAAGAAACGTTAAAAGAAGAAGAAAAACTAATTTAG 651

Qy   629 AAGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATA 688
              ||||| || || |||| | || ||| | ||| || ||||| || || |||
Db   652 ATTCTCTTCGAAACGAAAAAATATGCAGCAATCAGAACTTCTAATAGTTTTATGATGAA 711

Qy   689 AAGCTAGTAGTTTAATAACTAAAACACTAGATACACTACTTTAGGGGGAACGCTTTAGATT 748
              ||||| ||| ||| || || | | | | || |||| | | |||| |||| ||||
Db   712 AAGCCAATACTAATTGTTACACGACTACTTTAGACCCTGCTACTGGAAATATTCCTGAAGTTA 771

Qy   749 CTAATGAGATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAAC 808
              ||||| | |||| || |||| || ||| | | ||||||| ||| || ||| ||
Db   772 TGAGTG---TAACACAAGCTAATCAAGATATTAGCTAACAGTTTTATCAAGACTAATAGCTT 828

Qy   809 AAAAGACTAATGCTATGCTAAAAGATTATCTAATAGTTTTATTAAAAAAGTAATAAATAATG 868
              ||| |||| | || | | |||| | || ||||| |||| ||||| ||||
Db   829 GAAAACAAAATGCTGATAAATTTAGCTAACAGTTTTATCAAACAGTCTTTAGTTAAAAATA 888

Qy   869 AACAAAGTTTGTAGGGACTTTTACAAACGCTAATGTTCAACCTCAAACTACCAGTTTG 928

```
;  LENGTH: 3057
;  TYPE: DNA
;  ORGANISM: Ureaplasma urealyticum
US-09-601-198-55;

Query Match             5.0%;  Score 65.6;  DB 4;  Length 3057;
Best Local Similarity  46.2%;  Pred. No. 3.4e-05;
Matches 218;  Conservative  0;  Mismatches 254;  Indels  0;  Gaps  0;

Qy    444 AGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACACTAAAAATGGCTAAAAC 503
              |||  || |||| |||  ||||||| |||||  ||  | |||| |  |||||| ||||
Db   2350 ATCAGCTAAATTTAAAAAATAAGCCAGATTTAGCTAAATGAATGTTAACAATAAAATAAAAA 2291

Qy    504 TAATTTAGAATCAGCGCATCAACAAGCTAATACGGATAAAGCTTTTGATAATGAACA 563
              |||  | | ||||  | ||| || ||| || |||| |||  ||||| |||||||
Db   2290 TAATCTAATATATGGAGGTAATCTAGATACTAAATTCAAAACTAAATTAGTTGATAAAACTAAA 2231

Qy    564 CCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAA 623
              | |||||| | || |||  | | |||| | ||||||| || |||| | |  |||||
Db   2230 ATTATTGAAATTAGAATCAAATAATGATAATATTTATTACAATATACAAGAAAAATGCTGTTAA 2171

Qy    624 CCTTGAAGGTTTGTCATCAACTGCTTATATATAATCGCAATAATTAGTTGGATCATATA 683
              || | ||| ||| || |||| || || ||   | |||  |||  |  |  |||| ||
Db   2170 TCCTAATAGAGTGATAGGAACAATAAAAGCAGAATTTTCAAGAATTAATTCTAATTACAT 2111

Qy    684 CAATAAAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGGAACGCTTTT 743
              |  | |  | ||| || |||| ||  |    | |  || || ||| |  || |||
Db   2110 TAACAAAAAATAAATTAGTATATGAAGATAACAATCATCAAGAAATGAGAAACTAATGA 2051

Qy    744 AGATTCTAATGAGATTACTACAGTTAATCGGAATTATAATACGTTATCACTATTAA 803
              | |||  || |  || |||| ||| | |||| |||  |  | || | || |||
Db   2050 ACTAACTTTATCAAACAATAAAACTAGTTATAATTTGATATTAGTAATAAACTAACTCCTAA 1991

Qy    804 TGAACAAAGACTAATGCTGATGCATTAAGACTAATCTAATGAGTTTTATTTAAAAAAGTGATTCAAAA 863
              || | |  | | ||| | || |  || |||||  ||||  |||| || |||| |  | |
Db   1990 TAGAAAATATTCTTTCAAAAAGATTTAATTCAAATAATAATATCTTTGAAGAATTGCT 1931

Qy    864 TAATGAACAAAGTTTGTAGGGACTTTTACCAAACGCTAATGTTCAACCTTCA 915
              ||  |||||||  ||  | ||| ||| |||| | || || | ||||||||
Db   1930 AAATTCAAATAATAATATTGCTGATTCATTTATATATCATCATGCACCAACTGCA 1879
```

```
RESULT 8
US-10-204-708-73/c
;  Sequence 73, Application US/10204708
;  Patent No. 6677731
;  GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT: PIEPENBROCK, Christian
;  APPLICANT: BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION: by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/204,708
;  CURRENT FILING DATE:   2003-05-06
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE:  2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE:  2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE:  2000-09-01
;  NUMBER OF SEQ ID NOS: 98
SEQ ID NO 73
;  LENGTH: 5152
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73
```

```
Db   4685 CAATTCCACAAAACACCTTAAATAAATAAATAAAATAAAAACTAATAACCCAAAAACAAC 4626
              ||  || |||  |||| |||  ||  || |||||||||  || |||
Qy    175 TAAGATAAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATT 234
              | || |  |||||| |||||| ||||||||||  ||   |||||
Db   4625 AAACACTTTAATATATTCTACCCACACAACATTTTTATAAAATCCCAACAAACAATAAAA 4566
              |  |  ||| |||| |||  |  ||  ||| | | |||||||||
Qy    235 AGTTTGTTAGGTACACATCCTTCTTAGCATGGGATTTCTAGCTGTATGTCTTATTACT 294
              |||| | ||| || |  | | |||| ||||||  |||||| | | || |||
Db   4565 AACTACTTAAATAAATAAATAAATACAAAATCAAAATACAATATCCTAAAAATAAAA 4506
              || ||| |||| ||| ||  ||| |||   ||| | | ||||||||
Qy    295 AAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGGAGTTA 354
              |||||  ||||   ||| ||  || |  | |  || || |||  |||
Db   4505 AATAAAATAAAACAACAACTAAAATTCATTATTCTTTAATAAATTAACTTATA 4446
              ||   | |  |||  ||| | ||   ||  |||  ||||
Qy    355 ACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT 414
              ||  |||  || |||||  || |  || || ||  || ||| ||  ||||||
Db   4445 AAAAAATAAACAACATAAACAAAATATTTCTATTTAATAAATATCTAATATATCTAATA 4386
              ||   ||| |||| | | ||  |||| ||| | | || || ||||
Qy    415 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAGTTAAACCTTAATGCA 474
              ||| ||| ||| || ||| ||| |||| |||  | ||||  ||  ||||
Db   4385 AAATCCAAAAAATATTTTTTATATATAATAAAATTTAAAAAAAAACCTAATTATC 4326
              |||  | ||||||||  | | ||||||| ||||  ||| ||
Qy    475 ACATTAGAACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCAACCAAGCTAAT 534
              || ||| ||| || ||||  ||| || ||||  | | ||||  ||  |||
Db   4325 TAAACACTAAAAAAAAAAACTTCCTCTATATTCAACAATTAACCACAATATTTARA 4266
              || || |||||||| | ||| || ||| ||| |||||||||
Qy    535 ACGGATAAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTA 594
              | | |||| | | ||||| ||  ||||  |  |||||| | || |||  | ||||
Db   4265 ACATACTAAACACCCCAAAATTTACTAAAAAAAACAATTAACCAATAACAAA 4206
              | ||| ||  |||  |||| |||| |  | | ||||||
Qy    595 AAAAACCACTTTAGAACAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT 654
              |||||  |||||||| ||| | ||| | || ||| || ||||| ||||
Db   4205 AAAAACTAAAAATAACGAAAAAATACTAATAAAAATCATAAAATAATAAAAACC 4146
              |||||  |||||| ||  |  | |||| || ||||||  |||
Qy    655 CAAATTCGCAATAATTTAGTGGATCTATA--CAATAAAGCTAGTAGTTTTAATAACTAAAA 712
              ||||||  ||||  |||| || |||| |    |||  ||| | ||| | ||||
Db   4145 TCATATAAAAAAAATATACACGATATATATCTATTTAAAATAAAAAAACGACAA 4086
              || || |||||||||||| | |||| ||| ||||  |||||
Qy    713 CACTAGATCCACTAAAATGGGGGAACCGCTTTTAGATTCTAATGAGATTACTACAGTTAATC 772
              || ||| |||||||  | ||| ||||||| ||| ||  ||||||
Db   4085 AAATAAAAACATTAAAACTAAAACATTTTTTTATATATTACTCCTACCATCTAAAA 4026
              ||||||||  | | |||| ||  ||| | | |||||
Qy    773 GGAATTAATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTAT 832
              |||  ||||| ||||| ||||||  || ||||||||
Db   4025 TTTTTAATAAAATAAAATTATCTCAACTTTCAACTTTCAAACTAAAAAT-ATAATATATAAAACTT 3967
              | || |||| |||  || ||||||
Qy    833 CTAATAGTTTTTATTAAAAAGTGATTCAAAATAATGAACAA 874
              |  | ||||| ||| |||||| | |||
Db   3966 CAAAAAATTTATAACAACAATAAAAATAAATTACGAAA 3925
```

```
RESULT 7
US-09-601-198-55/c
;  Sequence 55, Application US/09601198
;  Patent No. 6531583
;  GENERAL INFORMATION:
;  APPLICANT: Cassell, Gail H.
;  APPLICANT: Chen, Ellson Y.
;  APPLICANT: Glass, Jennifer S.
;  APPLICANT: Glass, John I.
;  APPLICANT: Heiner, Cheryl R.
;  APPLICANT: Lefkowitz, Elliot
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;  TITLE OF INVENTION: UREALYTICUM
;  FILE REFERENCE: UAB-13452/22
;  CURRENT APPLICATION NUMBER: US/09/601,198
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 60/073,189
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 55
```

Qy   1044 TTCT 1047
Db   3532 TTCT 3529

RESULT 9
US-10-204-708-23//c
; Sequence 23, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERNCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 23
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-23

Query Match   4.5%; Score 59.4; DB 4; Length 11049;
Best Local Similarity 43.6%; Pred. No. 0.00092;
Matches 267; Conservative   0; Mismatches 346; Indels   0; Gaps   0;

Qy   1 AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Db   1317 AATAACTTAAAAATAATTAAAAAAATAAAAATAAAATATTAAAATAAAACGTTAACTC 1258

Qy   61 AATCCTAAATAAATAAGCCGTTAAATATCAATTAACTAAATAAAAATTAAAAATGGTTTTCTTATC 120

Db   1257 AAACCTCTTTCACTATCATCAAATTTCTAACTAACTATTTTACAAAATAATTTCATT 1198

Qy   121 AACCAAAATTCTCTAGTAATAAACCTTATTTATTTTTAGTCATCTTTAAGAT 180

Db   1197 ATCTTAAATACTTTACTAATTAACTAATTTAACTATAACGAATTCATCCTTACGCCT 1138

Qy   181 ATAAATATATCTTAAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Db   1137 TTATTTCACATTCGTTATTCAATATCAATATAAAAAAATTCTCCCAATAACCTACAAAACAAAAATC 1078

Qy   241 TTAGGTACAACATCCTTTCTTCCTAGCATTGGGATTTCTAGCGTGTATGTCTATTACTAAAAA 300

Db   1077 TAAATTAACTCTCACTTTCAAAACTTCTAAAACTTCAAAAACCGATTTAA 1018

Qy   301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGAGGTTAACAGAT 360

Db   1017 TTCTCTAATCCTTATAAATAAATAAAAAAACTCATAAAAAATAATTAATATATTAAAAACAC 958

Qy   361 CTAATCAATGCTAAAGCTAAAGCGATGCATTAGTTCACTACAAGACTTCACTCACCGT 420

Db   957 AATAAATACTTTATAAAATTTATAAAATCAATCAAAAAAACATATTATTTTAAAAACT 898

Qy   421 AGTTTATCATCGTCGTTATAGTGAAGCTGAAACAGTTGAAACATTTATATAT 480

Db   897 ATATTAACAACCACTAAAAAAAAACATAAAAAATAATTAATTAAAAAATCAACAAAACT 838

Query Match   4.7%; Score 61.2; DB 4; Length 5152;
Best Local Similarity 41.9%; Pred. No. 0.00033;
Matches 429; Conservative   0; Mismatches 593; Indels   2; Gaps   1;

Qy   24 ATATCTTTGCTTAAAAAAACACAAAATCTTCTACAAAAATCCTAAATAAATAAGCCGTTA 83

Db   4550 ATAATACTTCTAAAAATCCTAAATAAATAAATAATTATTTATCCTAAATACTAAAAACTAC 4491

Qy   84 AATTAACTAAAAAAAAATTAAAAAAATGTTTTCTTATCAACCAAAATTCTCTAGTAATAAA 143

Db   4490 CACATATTACTTTTTAAACCTTTTTTTTTTTTTTTTAACTATTTTATTAAAAAATAAA 4431

Qy   144 CGCTTATTTATTTTTATTTTTAGTCATCTTTAAGATATAAAATATATCTTAAATTCTAT 203

Db   4430 CCTAAAATAAAATTCAAAACACTTTTTAAATTAAAATAAAATCAAAACCTATTTCTAAATACCA- 4372

Qy   204 GAATAAGAAAAGAATCATCTTAAAGACTATTAGTTGTTAGGTAGGTACAACATCCTTTCTTAG 263

Db   4371 -AATTATTATTTATTTTTAAACAATAAAATATTTTACAACAAAAACATATACCGAA 4313

Qy   264 CATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCA 323

Db   4312 CATAATAACTCATACCTATAATCCCAACACTTTAAAAAAACCGAATAAATAAACACTTA 4253

Qy   324 AACCCAATTAGAAGCAGCGCGGAATGAGGTTAACAGATCTAATCAATGCTAAAGCCGATGAC 383

Db   4252 AATCCAAAAATTCCAAACCATCCTAAACAATAATAAAACCCTATCTTCACAAAAATAAA 4193

Qy   384 ATTAGCTTCACTCACAGACTATGCCAAGATGAAGCTAGTTTATCATCTGCTTATAGTGA 443

Db   4192 AAAAAAAAATAAAAAATAAAAATTAACCAAATAAAAAAAATCACTTACCTACA 4133

Qy   444 AGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAC 503

Db   4132 ATCCCAACTACTCAAATCTCAACAAAAATCTCAAACAAAAATTAAACCCTATCTCA 4073

Qy   504 TAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACA 563

Db   4072 TCATAAACCGTAATTACGCCACTACATTTCAAACCAAACCCTATCTCA 4013

Qy   564 CCCAAATTTAGTTGAAGCATACAAAGCACTTAGACAACAACGTGCTACTAA 623

Db   4012 AAAAAAAAAAAAAAAATTTAAAACAATAAATAACAACGCCATCATCATATA 3953

Qy   624 CCTTGAAGGTTTGTCATCAACTGCTTATAATCGCAATAATTTAGTGGATCTATA 683

Db   3952 CTATTATTATACTTATATTAAAAAAATACTTTAAAATTTTTAAATTTTTTATTT 3893

Qy   684 CAATAAAGCTAGTAGTTTAATAACTAGAACACTAGATCCCACTAAATGGGGAACGGCTTTT 743

Db   3892 ATCTTAAATACTTTACTAATTAAATTCAAAAAAATATATACAAAATTTATTACGACGAACAT 3833

Qy   744 AGATTCTAATGAGATTACTACAGTTAATCGGAATAATAATACGTTATCGATCAACCATTAA 803

Db   3832 AAAATCAATTTTTCTTTAACACCCTAAAATAAATTAAAAATAATTTCATTAAAAACTAA 3773

Qy   804 TGAACAAAAGACTAATGCTGATGCATTAATCTTTATTAAGTTTATTAAAAAGTGATTCAAAA 863

Db   3772 TTTCCTTTAACACAATCTTTAATTTATAAATATTTTCCCATTTTTAAAAAAATTCCTAA 3713

Qy   864 TAATGAACAAAGTTTTGTAGGGACTTTTACAAACGCTAATGTTCAAACCTTCAAACTACAG 923

Db   3712 AATACCTAAATATTTAAAAAACAACTACTTACACTAATTCTTACAAACTATAT 3653

Qy   924 TTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAATTATAAAATATGCAAGAAGACCGT 983

Db   3652 TTTACTTATTTTCTTTACATTTACAAAATCAAAATTTTTACAAAAAATCAGATTATTTCCAACTC 3593

Qy   984 TTGGAATGGTGATGAACCTTCAGTAGGATTCCTTGCAAACACGAATAGTATCACAGATGT 1043

Db   3592 TTATTTTAAAAATTAAATTAATAATAAAAATCTCCTTAAACTAAAAAATATAAAAACACTTTTATT 3533

```
Qy  481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
            ||  ||| |||| ||  ||||||
Db  837 TAAAAACAATAAATATATAAACCATCCATTATAAACTATTACCAAAAAAATTCCCGAT 778

Qy  541 AAAACGACTTTGATGAATGAACACCCAAATTAGTTGAAGCATACAAGCCACTAAAAACC 600
            ||||| || ||||| ||
Db  777 CCAAATACCGAAAAAATTCTTAAACCTTAAATAAAAAAAAATTCGAAACAAATCCATAAA 718

Qy  601 ACTTTAGAACAAC 613
            || ||||
Db  717 ATAAAATAAAAAC 705

RESULT 10
US-09-417-485D-5
; Sequence 5, Application US/09417485D
; Patent No. 6541202
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; FILE REFERENCE: 47714-5009-US
; CURRENT APPLICATION NUMBER: US/09/417,485D
; CURRENT FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 10640
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (834)..(7385)
; OTHER INFORMATION: TERT gene
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1821)..(1837)
; OTHER INFORMATION: m at position 1821 = a or c; w at position 1837 =
; OTHER INFORMATION: a or t. Xaa (amino acid) at position 330 = Leu or
; OTHER INFORMATION: Ile; Xaa at position 335 = Asp or Gly.
US-09-417-485D-5

Query Match             4.2%;  Score 55;  DB 4;  Length 10640;
Best Local Similarity   42.7%;  Pred. No. 0.0078;
Matches  336;  Conservative   0;  Mismatches 450;  Indels   1;  Gaps   1;

Qy  111 TTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTATTTATTTTATTTTTAGTCAT 170
Db 3102 TTTTGTCTTTACCATCAAATTTTAGGATTAAAGTTAAAAAAAGAATGAACCCAA 3161

Qy  171 CTTTTAAGATATAAATATATCTTAATATTCTAATAAGAATAAGAAAATCATCTTAAAGAC 230
Db 3162 ATGGATTATGTACAAAATATGTTCAATGTGAAGAAAAAGGTGAAAAATAAAAACAAT 3221

Qy  231 TATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTAT 290
Db 3222 AAATATATATTTATTAAGAAAATGAAAAAAAGAGCACTAATAAATGTATTAATAATAAA 3281

Qy  291 TACTAAAAAGATGCAAACAACCCAAATAATGGCCAAACCCAATTAGAGCGCGCGAATGGA 350
Db 3282 TTTTCAAAAAATGTATCCCTAAAAAAAAAAAATTTATATAACATAACATCACACGTCAT 3341

Qy  351 GTTAACAGATCTAATCAATGCTAAAGCGATGACTTAGCTTCACTACAAGACTATGCCAA 410
Db 3342 AATATATATTTATTAAAAGGATATGGAAAAAAATCAAAAACTAACAAATTTATTAAT 3401

Qy  411 GATTGGAAGCTAGTTATCATCGTCGTTATAGTGAAGCTAAACAGTAACAATAACCTTAA 470
Db 3402 AAAGTATAGATAATTTATACAAATTAAAG-GAAATTAACAAAAAAAGTGTTAGACCATA 3460

Qy  471 TGCAACATTAGAACAACTAAAAATGGCTAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGC 530
Db        | | |  |  |     |  |      | |    |
```

```
Db 3461 TATTAAAAAATTTACTATAAAAATAAAAAAAGAAATATTTGCTCTAAAAAAAATGTATAT 3520
            || ||||  |||||
Qy  531 TAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGC 590
            ||  ||| |
Db 3521 TCATATGAGAATGGCAAAAGAAGAGAAAAAAGTAACATAAAATTAGAAAAAGAGCATTCAAACA 3580
            |||||||  |
Qy  591 ACTAAAAACCACTTAGACAACACGTGCTACTAAACCTTGAAGGTTGTCATCAACTGCTTA 650
            || || |||||| ||  |
Db 3581 TTTTTTTATTTTGCTCAAGAAAAAGAACACATATTGAAATATTTTAGTTCCCATTTTT 3640
            |  || |
Qy  651 TAATCAAATTCGCAATAATTAGTGGATCTATACATAAAGCTAGTAGTTTAATAACTAA 710
            ||  | ||||  | | ||
Db 3641 TCAAAATAGAAAGATAAATTATGGTAAACGATTTAATAAACTAATCATCGAATAAAAAA 3700
            |
Qy  711 AACACTAGATCCACTAAAATGGGGGAACGCGTTTAGATTCTAATGAGATTACTACAGTTAA 770
            | ||||  | ||| ||
Db 3701 TATAATATAAAGCAAACAGTGGAATTGTTAAAAATAAGGATAAGACATTTTACATTT 3760
            |
Qy  771 TCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATT 830
            | ||| ||||  |
Db 3761 AATCAAAAATAAAAGTAACAAAAATAACAATAACAGAGAAGAACAAAAATAATAATAATAA 3820
            |||  || |||||
Qy  831 ATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTT 890
            |  |||  |||||
Db 3821 CAATAATAATATTAATAATAACAATAATAACAATAATAATAATAATAATAA 3880
            |
Qy  891 TACAAAC 897
            ||||||
Db 3881 TAACAAC 3887

RESULT 11
US-09-790-988-1
; Sequence 1, Application US/0790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match             4.1%;  Score 54;  DB 4;  Length 640681;
Best Local Similarity   56.9%;  Pred. No. 0.03;
Matches   99;  Conservative   0;  Mismatches  75;  Indels   0;  Gaps   0;

Qy     16 TTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAAAAATCCTAAATAAATA 75
Db 604204 TTTATTTAAAATGCTGTTTATAAAAAAAAAAAGAATAGCATTCTTTCAAATTTTGTAAAAATA 604263

Qy     76 AGCCGTTAAATTAACTAACTAAAAAATTAAAAAAATTAAAAAAATGTTTTTCTTATCAACCAAAATTCTCTA 135
Db 604264 AGTTTTATAAATAATTTTAAAATTCATCTAATTATTTTTTAATTAAATTATATTCTAAA 604323

Qy    136 GTAATAAACGCTTATTATTTTATTTTTAGTCATCTTTTTAAGATATAAATATA 189
Db 604324 ATAATAGAATAGACTTCTTTGCTATTTACAGGTAAATCACAAATTTTATTATA 604377

RESULT 12
US-10-204-708-82/c
```

```
Query Match              4.1%;  Score 53.6;  DB 4;  Length 1887;
Best Local Similarity   43.4%;  Pred. No. 0.011;
Matches 349;  Conservative  0;  Mismatches 449;  Indels  6;  Gaps  2;

Qy     82 TAAATTAACTAAAAAATTAAAAAAATGTTTTCTTATCAACCAAAATTCTCTAGTAATA  141
          ||||                    |||||      ||||||| |||||||  ||||| |
Db   1284 TTAGTTAATAATAACACAAAGCTAATTTAAAAAACTATATGAAGATTTAAATTAATT 1225

Qy    142 AAGCTTATTTATTTTATTTTTTAGTCATCTTTTAAGATACATTAATATATATAATTCT  201
          ||| |||||||||||    |||||||||| |||| ||||||||  ||||||||||||
Db   1224 TACACTAAATTAAATGATATTAATGATTTAATTAACAAGAACTAAAAGCTTTTACCTTTT 1165

Qy    202 ATGAATAGAAAAGAATCATCTTAAAGACTATTA----GTTTGTTAGGTACACATCCT  256
          |||       |||        |||||| ||||||
Db   1164 ATTAAAAAATTAGATGATATTGTTAATGTTATTATTACCAACAAAATTATAGT 1105

Qy    257 TTCTTAGCATTGGGATTTCTAGCTGTATGTCTTCATTACTAAAAAAGATGCAAACCCAAATA  316
          ||||          |||     ||||   |||||   ||  |||||  |||
Db   1104 TTTTTATGATTTGATCAATTAAAATTAACTAATTTAAACAAAATCTAAACAAGAATATC 1045

Qy    317 ATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAG  376
          |||     ||||
Db   1044 ACTAATATTGTTAATGTTAAAAAATCATTAGATATTGAACAAATTAAATCTTTAACTAAT 985

Qy    377 CGATGACATTAGCTTCACTACAGACTATGCCAAGATTGAA-GCTAGTTTATCATCTGCT  435
          |||       |||   ||||||
Db    984 AAATTATTAAATGATTTACAAAATTATTTACAAGGTTTAAATACAATTGCATTAGATAGT 925

Qy    436 TATAGTGAAGCTGAAACAGTTAACAATAACCTAATGCAACATTAGACAACTAAAAATG  495
          ||||||
Db    924 GAAAAATAGTATTAATAAAATAATTAGTATTAGTATTATTCAATCATTAAAACCAGAGATTACAAT 865

Qy    496 GCTAAAAACTAATTTGAATGCAGCCATCAACCAAGCTAATACCAAAAGACAAAAAACAATAGTTCAAAA  555
          |||
Db    864 CCTTACTATGTATTGAAAAACAACAAATATCCAAAAGCACTAAAAAACCACTTTAGAACACACGT 805

Qy    556 AATGAACAACCCAATTAGTTGAAGCATACAAAGCATACAAAGCAGCTAAAAAACCACTTTAGAACACACGT  615
          ||||||
Db    804 GCTACATAATAATGTTTTAGAAAATTACAATTAGAAAATTACAATTCAATTATTAGTTTATTT 745

Qy    616 GCTACTAACCTTGAAGGTTTGTCATCAACGATTGAGTAGTGCAATAATTCGCAATAATTAGTG  675
          ||||
Db    744 AGCTTAATACAAATAGCTAAGATTGTGCAAAATAATCTTAATAATGATGATTAACT 685

Qy    676 GATCTATACAAATAAAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGGA  735
          ||
Db    684 AAAAAAAATTTGAAGATGCTAATACTAAATTAGTTCAATTGTTGATCAAATCAAACAAACA 625

Qy    736 ACGGCTTTAGATTCTAATGAGATTACTACAGTTAATCCGAATATTAATAATAGTTATCA  795
          ||||||    |||
Db    624 ATTCTATCATCAATTTAAAAAAACTTATTTATTTATTCAAAGATTTCATTAAAGATTTCATTAAAAAGAA 565

Qy    796 ACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTCAATAGTTTTTATTAAAAAGTG  855
          ||||
Db    564 TTTAATACTAAAGTAATAGTGAAAATGATCAAGCAAAGAAAACTAATAAAACCCAAACA 505

Qy    856 ATTCAAAATAATGAACAAAGTTTT  879
          ||||||
Db    504 ATTCAAGATGTTTTAAATAGTTAT 481
```

```
; Sequence 82, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;   TITLE OF INVENTION: by Assessing DNA Methylation
;   FILE REFERENCE: 5013.1012
;   CURRENT APPLICATION NUMBER: US/10/204,708
;   CURRENT FILING DATE: 2003-05-06
;   PRIOR APPLICATION NUMBER: PCT/EP01/03971
;   PRIOR FILING DATE: 2001-04-06
;   PRIOR APPLICATION NUMBER: DE 10019058.8
;   PRIOR FILING DATE: 2000-04-06
;   PRIOR APPLICATION NUMBER: DE 10019173.8
;   PRIOR FILING DATE: 2000-04-07
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 82
;   LENGTH: 5476
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-82
```

```
Query Match              4.1%;  Score 53.8;  DB 4;  Length 5476;
Best Local Similarity   58.4%;  Pred. No. 0.012;
Matches 94;  Conservative  0;  Mismatches 67;  Indels  0;  Gaps  0;

Qy     52 TTCTAACAAATCCTAAATAAATAGGCCGTTAAATTAACTAAAAAATTAAAAAAATGGTT  111
          ||  ||    |||||  ||
Db   5470 TTATATTTAAAAAAAAAAAACTAAAAAACATTACATAAAACATTACCAAAATATACGCTA 5411

Qy    112 TTTCTTATCAACAACCAAATTCTCTAGTAATAAAACGCTTATTTTATTTTTTTTAGTCATC  171
          ||||||       ||||||
Db   5410 ACTTTTAAACCCCAATAAACATTTAAAACATTATCCTATTTTATTAATTTTCTTCATC 5351

Qy    172 TTTAAGATAATAAATATATCTTAATATTCTTAATATTCTTAATATTCTTAATATCATGAATAAGAA  212
          ||||||
Db   5350 ATTTATTATATAATAATCAATTTTTAAATTTTTTAAATTCCTATAATATAAAA 5310
```

```
RESULT 13
US-09-601-198-39/c
; Sequence 39, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;   APPLICANT: Cassell, Gail H.
;   APPLICANT: Chen, Elison Y.
;   APPLICANT: Glass, Jennifer S.
;   APPLICANT: Glass, John I.
;   APPLICANT: Heiner, Cheryl R.
;   APPLICANT: Lefkowitz, Elliot
;   TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;   TITLE OF INVENTION: UREALYTICUM
;   FILE REFERENCE: UAB-13452/22
;   CURRENT APPLICATION NUMBER: US/09/601,198
;   CURRENT FILING DATE: 2000-12-08
;   PRIOR APPLICATION NUMBER: 60/073,189
;   PRIOR FILING DATE: 1998-01-30
;   NUMBER OF SEQ ID NOS: 181
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 1887
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
US-09-601-198-39
```

```
RESULT 14
US-10-204-708-14/c
; Sequence 14, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;   TITLE OF INVENTION: by Assessing DNA Methylation
;   FILE REFERENCE: 5013.1012
;   CURRENT APPLICATION NUMBER: US/10/204,708
```

; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 14
; LENGTH: 6113
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-14

Query Match   4.1%;   Score 53.2; DB 4; Length 6113;
Best Local Similarity 46.3%;   Pred. No. 0.017;
Matches 253; Conservative 0; Mismatches 283; Indels 10; Gaps 2;

Qy   24 ATATCTTTGCTAAAAAAACACAAAAATCTTCCAACAAAATCCTAAATAAATAAGCCGTTA 83
             ||| ||| |||||||| |||| || |||||||| |||||||||||
Db   4095 AAATTATACATTTTAAACATAAACAATTTTTTTATCAATTATACTTCAATTAAAAACTCTC 4036

Qy   84 AATTAACTAAAAAATTAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAA 143
Db   4035 AAAATACAATTACAATTATATAAAAAAATATACAAATTAATAAAATATATATAAAAATAAA 3976

Qy   144 CGCTTATTTATTTTTATTTTTTAGTCATCTTTTAAGATATAAATATATCTTAATATTCTAT 203
Db   3975 ATATTCATATTATTATAAAAAATAAACATAACTATTAAATTTAAAACTTAATATATATAT 3916

Qy   204 GAATAAGAAAGAAATCATCTTAAAGACTTAAAGACATAGTTGTTAGGTACAACATCCTTCTTAG 263
Db   3915 CTCTATAAAATAATAACAAAAACTAAATAATAATAATATTA-------TTTTATTAT 3864

Qy   264 CATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGCGCA 323
Db   3863 TTTTAAACTTATTTTTTTTTTTTTTTAAAATAATAAAATCTCACTTCATCACCCAAACTA 3804

Qy   324 AACCCAATTAGAGAGCAGCGCGAATGGAGTTAACAGATCTAAATGCTAAAGGCGATGAC 383
Db   3803 AAATACAATAAATAATAAATCTCGACTCACTACAACATCCGCCTCCCAAATTCAAATAATTTT 3744

Qy   384 ATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGA 443
Db   3743 CCTACCTTAACCTCCCGACTAACTAACTAAATAACAAAC--GCGTACCACCATACCCAACTAA 3686

Qy   444 AGCTGAAACAGTTAACATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAC 503
Db   3685 TTTTTATATTTTTAATAAAAATAAAAATTTCACCATATTAACCAAACTAATCTAAAAACTCC 3626

Qy   504 TAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACA 563
Db   3625 TAATTCTCAAATAATCCACCCACCTCGACCTCCCAAAATAATAAATTACAAACATAAACC 3566

Qy   564 CCCAAA 569
Db   3565 ACCGAA 3560

RESULT 15
US-10-204-708-12/c
; Sequence 12, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 12
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-12

Query Match   4.1%;   Score 53.2; DB 4; Length 6317;
Best Local Similarity 49.3%;   Pred. No. 0.017;
Matches 139; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy   34 TTAAAAAACACAAAATCTCTAACAAAATCCTAAATAATAAATAAGCCGTTAAATAACTAA 93
Db   5042 TTTATCAAAAAAAAAAATACTTAAATAAATCTCCTTAACAAAAAAACCAACATTAACAATA 4983

Qy   94 AAAATTAAAAAAATGGTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTAATTA 153
Db   4982 AAAAATAAAAAAAAAAAAAAAAAATTACATTATAAAAAAACCAAAATCATAAAACTAAATCTA 4923

Qy   154 TTTTTATTTTAGTCATCTTTTAAGATATAAATAATTATATATCTATGAATAAGAAA 213
Db   4922 TTTTATCTTAACAAATATTATTATTATTAATAACTAATACTATACTACTTTTTAAAAAAATTCA 4863

Qy   214 AGAATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATT 273
Db   4862 TTTTTCATATATATTTTTCATACAAAATTATATAAAATTATTACTAATATTTTCACT 4803

Qy   274 TCTAGCTGTATGTCTATTACTAACTAAAAAAGAGATGCAAACCCAAAT 315
Db   4802 TTTAACTTTTACAAAATAATAAATAAATATTTTACAAAAAT 4761

Search completed: May 5, 2004, 20:04:08
Job time : 106.495 secs

```
            GenCore version 5.1.6
      Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May  5, 2004, 17:07:04 ; Search time 514.265 Seconds
                                    (without alignments)
                                    11489.234 Million cell updates/sec

Title:          US-09-901-572A-1
Perfect score:  1306
Sequence:       1 aaaaacatcagattgttaat.........taaaatcgtttatcaggtt 1306

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2937390 seqs, 2262062796 residues

Total number of hits satisfying chosen parameters:    5874780

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published_Applications_NA:*
         1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
         2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
         3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
         4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
         5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
         6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
         7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
         8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
         9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
        10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
        11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
        12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
        13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
        14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
        15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
        16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
        17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
        18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
        19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match %	Match Length	DB	ID	Description
1		1306	100.0	1306	10	US-09-901-572A-1	Sequence 1, Appli
2		1028	78.7	1152	15	US-10-131-591A-1	Sequence 1, Appli
3		1024.2	78.4	1371	9	US-09-147-052-1	Sequence 1, Appli
4		1024.2	78.4	3261	9	US-09-147-052-3	Sequence 3, Appli
5		921.6	70.6	1082	15	US-10-131-591A-24	Sequence 24, Appl
6	c	73.4	5.6	4985	15	US-10-094-240-10	Sequence 10, Appl
7	c	73.4	5.6	4985	15	US-10-056-405-10	Sequence 10, Appl
8	c	73	5.6	6292	13	US-10-221-714A-461	Sequence 461, App
9	c	69.6	5.3	7892	17	US-10-257-166-138	Sequence 138, App
10	c	68.6	5.3	8136	15	US-10-311-455-528	Sequence 528, App
11	c	68.6	5.3	8136	17	US-10-257-166-38	Sequence 38, Appl
12	c	68.4	5.2	11836	15	US-10-239-676-102	Sequence 102, App
13	c	68.4	5.2	11836	15	US-10-240-453-114	Sequence 114, App
14	c	68.2	5.2	8392	15	US-10-311-455-1463	Sequence 1463, Ap

		ID				Description
c	15	US-10-221-613-101	13	12405	5.2	68	Sequence 101, App
c	16	US-10-239-676-35	15	12405	5.2	68	Sequence 35, Appl
c	17	US-10-240-453-43	15	12405	5.2	68	Sequence 43, Appli
c	18	US-10-312-841-1	15	3673778	5.2	68	Sequence 1, Appli
c	19	US-10-257-166-36	17	8076	5.1	67	Sequence 36, Appl
c	20	US-10-204-708-6	15	6669	5.1	66.4	Sequence 6, Appli
c	21	US-10-311-455-166	15	6669	5.1	66.4	Sequence 166, App
c	22	US-10-240-589C-6	17	6669	5.1	66.4	Sequence 6, Appli
c	23	US-10-221-714A-333	13	6264	5.0	65.8	Sequence 333, App
c	24	US-10-240-452-67	15	6264	5.0	65.8	Sequence 67, Appl
c	25	US-10-349-680-148	15	3057	5.0	65.6	Sequence 148, App
c	26	US-10-311-455-1221	15	5163	5.0	65.6	Sequence 1221, Ap
c	27	US-10-311-455-1797	15	8771	5.0	65.4	Sequence 1797, Ap
c	28	US-10-311-455-1	15	7351	5.0	64.8	Sequence 1, Appli
c	29	US-10-311-455-950	15	18512	5.0	64.8	Sequence 950, App
c	30	US-10-312-841-2	15	3673778	5.0	64.8	Sequence 2, Appli
c	31	US-10-311-455-1999	15	17594	4.9	64.2	Sequence 1999, Ap
c	32	US-10-311-455-2130	15	16766	4.8	63.2	Sequence 2130, Ap
c	33	US-10-221-714A-114	13	9254	4.8	63	Sequence 114, App
c	34	US-10-239-676-86	15	9254	4.8	63	Sequence 86, Appl
c	35	US-10-240-453-92	15	9254	4.8	63	Sequence 92, Appl
c	36	US-10-311-455-326	15	5883	4.8	62.6	Sequence 326, App
c	37	US-10-240-485-26	15	5883	4.8	62.6	Sequence 26, Appl
c	38	US-10-221-714A-291	13	8591	4.8	62.6	Sequence 291, App
c	39	US-10-311-455-1111	15	5678	4.8	62.2	Sequence 1111, Ap
c	40	US-10-311-455-114	15	6061	4.8	62.2	Sequence 114, App
c	41	US-10-311-455-1876	15	9728	4.8	62.2	Sequence 1876, Ap
c	42	US-10-257-166-174	17	9728	4.8	62.2	Sequence 174, App
c	43	US-10-311-455-1492	15	5520	4.7	62	Sequence 1492, App
c	44	US-10-221-613-235	13	6012	4.7	61.4	Sequence 235, App
c	45	US-10-172-086-60	15	7001	4.7	61.4	Sequence 60, Appl

ALIGNMENTS

```
RESULT 1
US-09-901-572A-1
; Sequence 1, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1306
; TYPE: DNA
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 gene
US-09-901-572A-1

Query Match            100.0%;  Score 1306;  DB 10;  Length 1306;
Best Local Similarity  100.0%;  Pred. No. 7.1e-229;
Matches 1306;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 AAAAACATCAGATTGTTAATCGATATCTTGATATCTTTGCTTAAAAAAACACAAAAATCTTCTAACAA 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AAAAACATCAGATTGTTAATCGATATCTTGATATCTTTGCTTAAAAAAACACAAAAATCTTCTAACAA 60

Qy        61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATTAAAAAAAATGGTTTTCTTATC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATTAAAAAAAATGGTTTTCTTATC 120

Qy       121 AACCAAAATTCCTAGTAATAAACGCTTATTTTATTTTTATTTTTAGTCATCTTTAAGAT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AACCAAAATTCCTAGTAATAAACGCTTATTTTATTTTTATTTTTAGTCATCTTTAAGAT 180

Qy       181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
```

```
Db   1261 CCAGCAGTTGATGATGAGATTAAGAGTTGCTAAAAATCGTTTATCGAGGTT 1306


RESULT 2
US-10-131-591A-1
; Sequence 1, Application US/10131591A
; Publication No. US2003059799A1
; GENERAL INFORMATION:
;   APPLICANT: Nippon Zeon Co., Ltd.
;   TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
;   FILE REFERENCE: J209
;   CURRENT APPLICATION NUMBER: US/10/131,591A
;   CURRENT FILING DATE:  2002-08-15
;   NUMBER OF SEQ ID NOS: 79
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1152
;   TYPE: DNA
;   ORGANISM: Mycoplasma gallisepticum
;   FEATURE:
;   OTHER INFORMATION: TTM-1 gene (after EcoRI)
US-10-131-591A-1

  Query Match          78.7%;  Score 1028;  DB 15;  Length 1152;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-178;
  Matches 1028;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy  279 CTGTATGTCATTACTAAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAGC 338
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6 CTGTATGTCATTACTAAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAGC 65

Qy  339 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCTAAAAGCTTACTTCACTACA 398
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCTAAAAGCTTACTTCACTACA 125

Qy  399 AGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAA 458
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126 AGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAA 185

Qy  459 CAATAACCTTAATGCAACATTAGAACAACTAAAAATAGGCTAAAAATTAGAATCAGC 518
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  186 CAATAACCTTAATGCAACATTAGAACAACTAAAAATAGGCTAAAAATTAGAATCAGC 245

Qy  519 CATCAACCAAGCTAATACGGATAAATGAACAGACTTTGATAATGAACACCCAAATTTAGTTGA 578
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  246 CATCAACCAAGCTAATACGGATAAATGAACAGACTTTGATAATGAACACCCAAATTTAGTTGA 305

Qy  579 AGCATACAAAGACTAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTC 638
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  306 AGCATACAAAGACTAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTC 365

Qy  639 ATCAACTGCTTATAAATCGCAATAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAG 698
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  366 ATCAACTGCTTATAAATCGCAATAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAG 425

Qy  699 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACAAACACGAAG 758
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  426 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACAAACACGAAG 485

Qy  759 TACTACAGTTAATCGGAAATATTAATAATACGTTAAAAAAGTGATTCAAAATAATGAGAAT 818
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  486 TACTACAGTTAATCGGAAATATTAATAATACGTTAAAAAAGTGATTCAAAATAATGAGAAT 545

Qy  819 TGGTGATGCATTATCTATAGTTTATTAAAAAGTTTATTAAAAAGTTCAAATAATGAAAGTT 878
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  546 TGGTGATGCATTATCTATAGTTTATTAAAAAGTTTATTAAAAAGTTCAAATAATGAAAGTT 605

Qy  879 TGTAGGGACTTTTACAAACGCTAATGTTCAACCTACAGCTAATCCAACT 938
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  606 TGTAGGGACTTTTACAAACGCTAATGTTCAACCTACAGCTAATCCAACT 665

Qy  939 TGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTCGATGA 998
```

```
Db   181 ATAAATATATCTTAATATTCTATGAATAGAGAAAGAATCATCTTAAAGACTATTAGTTTG 240
Qy   241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
Db   241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
Qy   301 GATGCAAACCCAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGAT 360
Db   301 GATGCAAACCCAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGAT 360
Qy   361 CTAATCAATGCTAAAGCTAAAAGCTTACTTCACTACAAGACTATGCCAAGATTGAAGCT 420
Db   361 CTAATCAATGCTAAAGCTAAAAGCTTACTTCACTACAAGACTATGCCAAGATTGAAGCT 420
Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
Db   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
Qy   481 GAACAACTAAAAATGGCTAAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
Db   481 GAACAACTAAAAATGGCTAAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
Qy   541 AAACGACTTTGATAATGAACAACACCCAAATTTAGTTGAAGCATACAAAGACACTAAAAACC 600
Db   541 AAACGACTTTGATAATGAACAACACCCAAATTTAGTTGAAGCATACAAAGACACTAAAAACC 600
Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
Qy   661 CGCAATAATTAGTGGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
Db   661 CGCAATAATTAGTGGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
Qy   721 CCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
Db   721 CCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
Qy   781 AATAATAACGTTATCAACTATTAATGATGCATTATCTATAAGT 840
Db   781 AATAATAACGTTATCAACTATTAATGATGCATTATCTATAAGT 840
Qy   841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTTAGGGACTTTTACAAACGCT 900
Db   841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTTAGGGACTTTTACAAACGCT 900
Qy   901 AATGTTCAACCTTCAAACTACAGTTTGTGTTGCTGTATGTAACACCCGTCAAT 960
Db   901 AATGTTCAACCTTCAAACTACAGTTTGTGTTGCTGTATGTAACACCCGTCAAT 960
Qy   961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
Db   961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
Qy  1021 AACACGAATAGTATCACAGATGTTCTTGGATTGTTCTTGGATTTATAGTTAGCTGAACAAACACGAAG 1080
Db  1021 AACACGAATAGTATCACAGATGTTCTTGGATTGTTCTTGGATTTATAGTTAGCTGAACAAACACGAAG 1080
Qy  1081 TACCAATTTAGTTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
Db  1081 TACCAATTTAGTTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
Qy  1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAATAATGAAAT 1200
Db  1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAATAATGAAAT 1200
Qy  1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTGGTGCAAATAATACTACAGCTAATCCAACT 1260
Db  1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTGGTGCAAATAATACTACAGCTAATCCAACT 1260
Qy  1261 CCAGCAGTTGATGATGAGATTAAAGTTGCTAAAAATCGTTTATCGAGGTT 1306
```

```
Db    666  TGCTGATGTGAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGA  725
                                                                       ||||
Db    428  CCATCAACCAGCTAATACGGATAAAACGACTTTGATAATGAACACCAAATTTAGTTGTTG    487

Qy    999  ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCCACAGATGTTCTTGGATTTATAG   1058
                                                                       ||||
Qy    578  AAGCATACAAAGCACTAAAAACCACTTAGAACACGTGCTACTAACCTTGAAGGTTTGT     637

Db    726  ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAG    785
                                                                       ||||
Db    488  AAGCATACAAAGCACTAAAAACCACTTAGAACACGTGCTACTAACCTTGAAGGTTTGT     547

Qy   1059  TTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGG   1118
                                                                       ||||
Qy    638  CATCAACTGCTTATAATCAAATTCGCAATAATAATTCGGATCTATACAATAAAGCTAGTA    697

Db    786  TTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGG    845
                                                                       ||||
Db    548  CATCAACTGCTTATAATCAAATTCGCAATAATAATTCGGATCTATACAATAAAGCTAGTA    607

Qy   1119  TTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGTGCTAATAACGTTGGATTACA  1178
                                                                       ||||
Qy    698  GTTTAATAACTAAAACACTAGATCCACTAAAATGGGGGAACGCTTTTAGATTCTAATGAGA   757

Db    846  TTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGTGCTAATAACGTTGGATTACA   905
                                                                       ||||
Db    608  GTTTAATAACTAAAACACTAGATCCACTAAAATGGGGGAACGCTTTTAGATTCTAATGAGA   667

Qy   1179  ATACAAATTAAATAATGGAAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAAA       1238
                                                                       ||||
Qy    758  TTACTACAGTTAATCGGAATATTAATATACGTTATCAACTATTAATGAACAAAAGACTA    817

Db    906  ATACAAATTAAATAATGGAAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAAA        965
                                                                       ||||
Db    668  TTACTACAGCTAATAAGAATATTAATATACGTTATCAACTATTAATGAACAAAAGACTA    727

Qy   1239  TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTT   1298
                                                                       ||||
Qy    818  ATGCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAATAATGAACAAAGTT     877

Db    966  TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTT   1025
                                                                       ||||
Db    728  ATGCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAATAATGAACAAAGTT     787

Qy   1299  ATCAGGTT   1306
                     ||||
Qy    878  TTGTAGGGACTTTACAAACGCTAATGTTCAAACCTTCAAACTACACAGTTTTGTTGCTTTTA  937

Db   1026  ATCAGGTT   1033
                     ||||
Db    788  TTGTAGGGACTTTACAAACGCTAATGTTCAAACCTTCAAACTACACAGTTTTGTTGCTTTTA  847

Qy    938  GTGCTGATGTAACACCCGTCAATTATAAATATAAAATATGCAAGAAGGACCGTTTGGAATGGTGATG  997

Db    848  GTGCTGATGTAACACCCGTCAATTATAAATATAAAATATGCAAGAAGGACCGTTTGGAATGGTGATG  907

Qy    998  AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATA   1057

Db    908  AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATA    967

Qy   1058  GTTTAGCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATGGTCCATCATCAACTG  1117

Db    968  GTTTAGCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATGGTCCATCATCAACTG  1027

Qy   1118  GTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTAC   1177

Db   1028  GTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTAC   1087

Qy   1178  AATACAAATTAAATAATGGAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAA       1237

Db   1088  AATACAAATTAAATAATGGAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAA        1147

Qy   1238  ATAATACTACAGCTAATCCAACTCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT  1297

Db   1148  ATAATACTACAGCTAATCCAACTCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT  1207

Qy   1298  TATCAGGTT   1306

Db   1208  TATCAGGTT   1216
```

RESULT 3
US-09-147-052-1
; Sequence 1, Application US/09147052
; Patent No. US2001001435A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US2001001435Aloru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: hybrid
US-09-147-052-1

```
Query Match      78.4%;      Score 1024.2; DB 9; Length 1371;
Best Local Similarity  99.7%; Pred. No. 2.2e-177;
Matches 1026; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

RESULT 4
US-09-147-052-3
; Sequence 3, Application US/09147052
; Patent No. US2001001435A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US2001001435Aloru
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147,052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28

```
Qy    278  GCTGTATGTCTAATTACTAACAAAAGATGCAAACCCAAATAATGCCAAACCCAATTAGAAG   337

Db    188  GCTGTATGTCTATTACTACTAAAAAAGATGCAAACCCAAATAATGGCAAACCCAATTAGAAG  247

Qy    338  CAGCGCGAATGGAGTGACTAACAGATCTAATCAATCTAAAGCGATGATGACATTAGCTTCACTAC  397

Db    248  CAGCGCGAATGGAGTGACTAACAGATCTAATCAATCTAAAGCGATGATGACATTAGCTTCACTAC  307

Qy    398  AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA   457

Db    308  AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA   367

Qy    458  ACAATAACCTTAATGCAACATTAGAACACTAAAAATGCTAAAACTAATTTAGAATCAG     517

Db    368  ACAATAACCTTAATGCAACATTAGAACACTAAAAATGCTAAAACTAATTTAGAATCAG     427

Qy    518  CCATCAACCAGCTAATACGGATAAAACGACTTTGATAATGAACACCAAATTTAGTTG      577
```

```
;   NUMBER OF SEQ ID NOS:  9
;   SOFTWARE:  PatentIn Ver. 2.1
;   SEQ ID NO 3
;   LENGTH: 3261
;   TYPE: DNA
;   ORGANISM: hybrid
US-09-147-052-3

Query Match         78.4%;  Score 1024.2;  DB 9;  Length 3261;
Best Local Similarity  99.7%;  Pred. No. 3.1e-177;
Matches 1026;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy   278  GCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG  337
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2078  GCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG  2137

Qy   338  CAGCCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC  397
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2138  CAGCCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC  2197

Qy   398  AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA  457
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2198  AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA  2257

Qy   458  ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAG  517
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2258  ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAG  2317

Qy   518  CCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG  577
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2318  CCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG  2377

Qy   578  AAGCATACAAAGCACTAAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT  637
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2378  AAGCATACAAAGCACTAAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT  2437

Qy   638  CATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAAGTAGTA  697
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2438  CATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAAGTAGTA  2497

Qy   698  GTTTAATAACTAAAACCTAGATCCACTAAATGGGGAACGCCTTTTAGATTCTAATGAGA  757
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  2498  GTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCCTTTTAGATTCTAATGAGA  2557

Qy   758  TTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA  817
          ||||| ||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  2558  TTACTACAGCTAATAAGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA  2617

Qy   818  ATGCTGATGCATTATCATTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTT  877
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2618  ATGCTGATGCATTATCATTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTT  2677

Qy   878  TTGTAGGGACTTTTACACACCCGTCAATTATAAATAATGCAAGAAGGACCGTTTGGATGG  937
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2678  TTGTAGGGACTTTTACACACCCGTCAATTATAAATAATGCAAGAAGGACCGTTTGGATGG  2737

Qy   938  GTGCTGATGTAACACCCCGTCAATTATAAATACGAATAGTATCACAGATGTTCTTGGATTAT  997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2738  GTGCTGATGTAACACACCCGTCAATTATAAATACGAATAGTATCACAGATGTTCTTGGATTAT  2797

Qy   998  AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCCACAGATGTTCTTTGGATTTATA  1057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2798  AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCCACAGATGTTCTTTGGATTTATA  2857

Qy  1058  GTTTAGCTGGAACAAAACGAAGTACCCAATTTAGTTTTAGCAACATAATGGTCCATCAACTG  1117
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  2858  GTTTAGCTGGAACAAAACGAAGTACCCAATTTAGTTTTAGCAACATAATGGTCCATCAACTG  2917

Qy  1118  GTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTAC  1177
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2918  GTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTAC  2977

Qy  1178  AATACAAATTAAAATAATGGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA
```

```
Db  2978  AATACAAATTAAAATAATGGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA  3037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  1238  ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT  1297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3038  ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT  3097

Qy  1298  TATCAGGTT  1306
          |||||||||
Db  3098  TATCAGGTT  3106

RESULT 5
US-10-131-591A-24
; Sequence 24, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE:  2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
;   LENGTH: 1082
;   TYPE: DNA
;   ORGANISM: Mycoplasma gallisepticum
;   FEATURE:
;   OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of
;   OTHER INFORMATION: pN240K-S
US-10-131-591A-24

Query Match         70.6%;  Score 921.6;  DB 15;  Length 1082;
Best Local Similarity  98.0%;  Pred. No. 1.1e-158;
Matches 933;  Conservative  0;  Mismatches  19;  Indels  0;  Gaps  0;

Qy   355  ACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT  414
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT  60

Qy   415  GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA  474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAACCTGCAGCA  120

Qy   475  ACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT  534
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   121  ACATTAGAACAACTACAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT  180

Qy   535  ACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA  594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA  240

Qy   595  AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT  654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT  300

Qy   655  CAAATTCGCAATAATTAGTGGATCTATACAGTAGTTTAATAACTAGTAGTTAATAACTAAAACA  714
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAAATTCGCAATAATTAGTGGATCTATACAGTAGTTTAATAACTAGTAGTTAATAACTAAAACA  360

Qy   715  CTAGATCCACTAAATGGGGAACGCCTTTTAGATTCTAATGAGATTACTACAGTTAATCGG  774
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CTAGATCCACTAAATGGGGAACGCCTTTTAGATTCTAATGAGATTACTACAGTTAATAAG  420

Qy   775  AATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCT  834
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AATATTCCAGAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCT  480

Qy   835  AATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA  894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA  540

Qy  1178  AATACAAATTAAAATAATGGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA  1237
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Qy 895 AACGCTAATGTTCAACCTCAAACTACAGTTTTGTTGCTGCTTTAGTGCTGATGTAACACCC 954
Db 541 AACGCTAATGTTCAACCTAGCCAGTAGTTTCAACCTAGCCAGTAGTTTGCTTGCTTTAGTGCTGATGTAACACCC 600
Qy 955 GTCAATTATAATAAATGCAAGAAGGACCCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT 1014
Db 601 GTCAATTATAATAAATGCAAGAAGGACCCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT 660
Qy 1015 CTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAAC 1074
Db 661 CTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAAC 720
Qy 1075 ACGAAGTACCAATTTAGTTTAGCAACTATGTCCATCAACTGGTTATTTATTATTTCCCT 1134
Db 721 ACGAAGTACCAATTTAGTTTAGCAACCATGGTCCATCAACTGGTTATTTATTATTTCCCT 780
Qy 1135 TATAAGTTGGTTAAAGCAGCTGGTCTGATGCTAATAACGTTGGATTACAATACAATCAAATTAAATAAT 1194
Db 781 TATAAGTTGGTTAAAGCAGCTGGTCTGATGCTAATAACGTTGGATTACAATACAATCAAATTAAATAAT 840
Qy 1195 GGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAAT 1254
Db 841 GGAAATGTTCAACCAGTTGAGTTTGCCACTTCAACTAGCGCACAGCAGCAGACTACAGCTAAT 900
Qy 1255 CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
Db 901 CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 952

RESULT 6
US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US2003008263TA1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N8289
; CURRENT APPLICATION NUMBER: US/10/094,240
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 10/056,405
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match   5.6%;   Score 73.4;   DB 15;   Length 4985;
Best Local Similarity   43.3%;   Pred. No. 0.0018;
Matches 392; Conservative 0; Mismatches 511; Indels 2; Gaps 1;

Qy 110 TTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCA 169
Db 3163 TATTGTGTATTACTACAGTTACTGGAATTATTTTACTATTTATTTACTTTTATTA 3104
Qy 170 TCTTTTAAGATATAAAATATATCTTAATATTCTTAATATTCTTATGATGAAGAAAAGAATCATCTTAAAGA 229
Db 3103 TTTTTATTATTATGATTACTGGAATAATTAGAGAATAAAAAACAATAATAATAAACAA 3044
Qy 230 CTATTAGTTGTTGGTACAACATCCTTCTTAGCATTGGATTCTAGCTGTATGTCTA 289
Db 3043 TAATAAGAATAACAACAACAACAATAATAAGAACAACAACAATAATAAGAATAA 2984
Qy 290 TTACTAAAAGATGCAAACCCAAATAATGGCCAAACCCATTAGAAGCAGCGCGAATGG 349
Db 2983 TAATAATAACAATAGCAATAATAAGCAATAATAATAACAATAGCAATAATAATAATAA 2924
Qy 350 AGTTAACAGATCTAATCAATGCTAAAGCGATGCATTAGCTTCACTACAACGACTATGCCA 409

Db 2923 TAATAATAATAATATAATAATAAAAAATAATAAATAATTATATAAGATAGTA 2864
Qy 410 AGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTA 469
Db 2863 ATAATAATAATAATAATAATAAAAATAATTAATAATAATAATAATAATAATAATAATA 2804
Qy 470 ATGCAACATTGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGCCATCAACCAAG 529
Db 2803 AGAATAATAATATATAATAATAATAATAATAATCATAATAAAA 2744
Qy 530 CTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAG 589
Db 2743 TTATTATTATGATAATAATGTATAATAATAACAATAATAA--TAATTAATAATAAT 2686
Qy 590 CACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTT 649
Db 2685 AAATTAATTAATTAATAATAATTATTGTTATAATAATAATAATA 2626
Qy 650 ATAATCAAATTCGCCAATAATTTAGTGGATCATACAATAAAGCTAGTAGTTTAATAACTA 709
Db 2625 ATAATAATAATAATCATAATTATCATAATAATAATAATAATA 2566
Qy 710 AAACACTAGATCCACTAAATGGGGGAACGCCTTTAGATTCTAATGAGATTACTACAGTTA 769
Db 2565 ATAATAATAATAATAATAATAATAAAAGTAATAATTATAATAATAATAGTA 2506
Qy 770 ATCGGAATATAATAATACGTTATCAACTATTAATAATAATTAATAAGACTAAATGCTGATGCAT 829
Db 2505 ATAATAATAATAATAATAATTATTATATAATAATAAT 2446
Qy 830 TATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGACAACAAGTTTGTAGGGACTT 889
Db 2445 ATAATAATTATAATAATAATAATAATAATAATAATAATAATAATAATA 2386
Qy 890 TTACAAACGCAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGTGCTGATGTAA 949
Db 2385 ATAATAATAATAATAATAATAATAATAATAAATTCTGTATTTTGGTTCCTGTAA 2326
Qy 950 CACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTA 1009
Db 2325 CTAAATTGGTTACTAAAAATGGTTTATCTTTCTCTTAAAATAAAAATCAATCCTAAAGGA 2266
Qy 1010 GAATT 1014
Db 2265 AGCTT 2261

RESULT 7
US-10-056-405-10/c
; Sequence 10, Application US/10056405
; Publication No. US2003016601A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: N7841
; CURRENT APPLICATION NUMBER: US/10/056,405
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-056-405-10

Query Match   5.6%;   Score 73.4;   DB 15;   Length 4985;
Best Local Similarity   43.3%;   Pred. No. 0.0018;
Matches 392; Conservative 0; Mismatches 511; Indels 2; Gaps 1;

Qy 110 TTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCA 169

RESULT 8

US-10-221-714A-461/c
; Sequence 461, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with

; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 461
; LENGTH: 6292
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-461

Query Match   5.6%;   Score 73; DB 13; Length 6292;
Best Local Similarity 43.9%; Pred. No. 0.0023;
Matches 368; Conservative 0; Mismatches 465; Indels 6; Gaps 1;

```
Qy   455 TTAACAATAACCTTAATGCAACATTAGAACAACTAACTAAAAATGGCTAAAACTAATTTAGAAT 514
             ||||  |  || ||||| ||| | |  |    |||    ||   || | |||||||
Db   466 AAAAATAAAACCAATAAATCAATAAATAAAAAAATAAAATAAAATAACGTAATAAAACATTAAATTCC 407

Qy   515 CAGCCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAG 574
             ||||| |||  | |||     |||  |||| ||  |    |||   | | |||||
Db   406 AACCTTACATCAATACATAAAATATCTCAATATTCTACTATAAAAAAATARAAAATTC 347

Qy   575 TTGAAGCATACAAAGCACTAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTT 634
             || ||| | ||||| |  |||  | |||    |||| |  |  ||||  |||||
Db   346 TTCAACCTAAATAATAAAAAAATCTAATTTATCATAAATATTAAATATAATAAAATC 287

Qy   635 TGTCATCAACTGCTTATATATCAAATTCGCAATAATAAAATTCTCAATAATACTACACTATTATCAAAGCTA 694
             |||  | | |||| ||||||  ||||    | || | |  | ||||    | | |||  ||||
Db   286 TTCAACTTCTAATTAATTAAATAATAATAATAAAATTCTCAATAATACTACACTTATTATCAAA 227

Qy   695 GTAGTTTAATAA 706
             | ||||| ||
Db   226 TAAATATAAAAA 215


RESULT 9
US-10-257-166-138/c
; Sequence 138, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
;   TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
;   FILE REFERENCE: 5013.1011
;   CURRENT APPLICATION NUMBER: US/10/257,166
;   CURRENT FILING DATE: 2002-10-07
;   PRIOR APPLICATION NUMBER: PCT/EP01/07470
;       DE 10032529.7
;       DE 10043826.1
;   PRIOR FILING DATE: 2001-06-29
;       2000-06-30
;       2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 138
; LENGTH: 7892
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-138

Query Match             5.3%;  Score 69.6;  DB 17;  Length 7892;
Best Local Similarity   46.0%;  Pred. No. 0.011;
Matches  309;  Conservative    0;  Mismatches  359;  Indels    4;  Gaps   2;

Qy    36 AAAAAAACACAAAATCTTCTACAAAAATCCTAAATAAATAAATAAGCCGTTAAATTAACTAAAA 95
             ||  |  || |||||| ||||||||||| ||||||||  ||    |||   | |||  |||
Db   883 ACAATAAATAAAAAAATACTTCATTAAATAAAAAAAAAAAAATAATCCAAAAAATAACTTTTA 824

Qy    96 AATTAAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACCGTCTTATTTATT 155
             ||| ||||||||             |||||||||| ||||||| ||  ||||  | ||
Db   823 AATATAAAAAAAAAAACAAAAAACAAAAAAAATTAATAAATATTTTTAAACTAAATAAATAAATA 764

Qy   156 TTTATTTTAGTGCATCTTTTAGATATAATATATCTTAAATATTCTATGAATAAGAAAAG 215
             || |||     | ||||||||  |||||||  ||||||||  ||
Db   763 TCAACTCATAAAAAAATAATTATTCCCAAAAATAATATACAAAATAAAATTAAAAT 704

Qy   216 AATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTC 275
             | |  |  |||| |||||||   |||||     || ||||
Db   703 ACATAACAATAACAATAACAATTAAAAAATTAAAAAATAAAATAAAATTAAA---AACTTAAAAAATTCT 647

Qy   276 TAGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCAAACCCAATTAGA 335
             |||||||||| |  ||||| |||||||  ||  |||   |||||||| ||||||||
Db   646 AAAATCATTACATTATCCTAAAAAAAAAAAAATACAATAATTAATCAATAATACATATC 587

Qy   336 AGCAGCCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACT 395
             | || |  | || |  ||||| | ||  ||| || |    || ||| | ||||  | |
Db   586 ATAATCCCTAAAAAAAAAATCATTATTAAATAAAAAAAAAAAAATAAAAATACATCTTATAAAA 527

Qy   396 ACAAGACTATGCCAAGAT-TGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAG 454
             ||||  | ||||  | || ||||    |||  | ||  |  || || |||| |||||
Db   526 AAAAATCTAAATAATAATATTAAACAAATCTAAATAAAATAAAAAACAAAATAAAAATAAAAA 467
```


RESULT 10
US-10-311-455-528/c
; Sequence 528, Application US/1031455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;   TITLE OF INVENTION: cytosine methylation
;   FILE REFERENCE: 5013.1014
;   CURRENT APPLICATION NUMBER: US/10/311,455
;   CURRENT FILING DATE: 2002-12-16
;   PRIOR APPLICATION NUMBER: PCT/EP01/07537
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 528
; LENGTH: 8136
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-528

```
Query Match             5.3%;  Score 68.6;  DB 15;  Length 8136;
Best Local Similarity   44.1%;  Pred. No. 0.016;
Matches  375;  Conservative    0;  Mismatches  471;  Indels    4;  Gaps   2;

Qy     2 AAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAAA 61
             |||| | ||  |||||| |||||  |     ||        ||  ||||
Db  1591 AAAATACTATTTTAATTAATTACTAACTACAACATAATAAACCCCAAAAATAAACCTTAACAT 1532

Qy    62 ATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATCA 121
             |  |||||||| ||    || || ||    |||||  |||||||||
Db  1531 TTAATAACTATTTAAAAATCAAAAAAATATATAAAAATATTAAAATTTTTTCTCAAATAA 1472

Qy   122 ACCAAAATTCTCTAGTAATAAACCGTCTTATTTATTTTTAGTCATCTTTTAAGATA 181
             ||||||| |||||| |||  |||  | ||  ||
Db  1471 AAAATAATTTAACATATTAACATAATATAAACACCCCGTTATATAATAATTACCA 1412

Qy   182 TAAATATATCTTAATATTCTATGAATAAGAAAATCATCTTAAAGACTATTAGTTTGT 241
             ||||||  |||||  ||||||||  ||    |  |  |||| |||||||
Db  1411 TATACAAAATAAAAACTACAATTAATTAAATAATTATAAACATAATATATATAT 1352

Qy   242 TAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAAG 301
             |||     ||  ||||          || |  |||||||||| |  |||||
Db  1351 TAAATCAAAAAAAATAACGCTATAATTATATAAAACTACTCCTACTAAATTAATAATAAAAA 1292
```

Qy   2 AAAACATCAGATTGTTAATCTGATATCTTTGCTAAAAAAACACAAAATCTTCTAACAAAG 61
        ||||| | ||| |||||||| || |||||||||||||||||| |||||||||| ||
Db 1591 AAAATACTATTTTAATTATACTATAACTACATAACAATATATATAACCCCAAATAACTTTAACAT 1532

Qy   62 ATCCTAAATAAATAAGCCGTTAAATTAACTAGAAAAAAATTAAAAAAAATGGTTTTTCTTATCA 121
        |||||||||||| |||||| |||||||||||||||||||||||||||||| |||||
Db 1531 TTAATAATAATAAAAAAATCAAAAAAATATAAAATATTAAAAATATTTTTCTCAAATAA 1472

Qy 122 ACCAAAATCTCTAGTAATAAACGCTTATTAATTTATTTTTAGTCATCTTTTAAGATA 181
        ||| |||||| ||||||||| || |||||||||| |||||||||||||||||||
Db 1471 AAATAATTTAACAATATTAATTATATATAAAACACCCCTCGTTATATAATTATTACCA 1412

Qy 182 TAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTGT 241
        |||||| || |||| |||||||| ||| ||||||| |||| |||| || ||||| |
Db 1411 TATCAAAAATAAAAACTACAATTATTAATCTTATAAACATAATATATAATAAT 1352

Qy 242 TAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTTCATTACCAAAAAAG 301
        |||||||||||||||||||| |||||||||||||||||||||||| ||||||||||||
Db 1351 TAAATCAAAAAATAAACGCTATATAATTATAAATCACTACTCCTACTAAATTAATAATAAAAA 1292

Qy 302 ATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATC 361
        ||||||||| |||||||| |||| ||||||||| |||||| | ||| ||||||| |||
Db 1291 ACAAAACAAAATTAACAAAACTACTAAAAATCATCACACAATAAATCAAAAAAAAAATA 1232

Qy 362 TAATCAATGCTAAAGCGATGACATTAGCTTCACTACACAAGATCATGCCAAGATTGAAGCTA 421
        |||||||||| |||||||||||||||| | |||||||||| ||| | |||||||||||||
Db 1231 TTTAAAACCTTATATTTATATATTAATATACTAAATTTTAAAAATACTACTATAAAACTA 1172

Qy 422 GTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAG 481
        |||||||| ||| |||||||| || || |||| |||||| |||| ||||||| ||
Db 1171 CAAAACAAAATAATTAATAAAAAATAATCAATTTAGAATCAGCCATCAACCAAGCTAATACGGATA 1112

Qy 482 AACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATA 541
        |||||||||||||| ||| ||| || ||||||||||| ||| |||||| ||||||||
Db 1111 TAAAACTTCAAAAATTTAAA-TAAAAAAACATAAAATTATTAATAAACATATACTTATA 1053

Qy 542 AAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGGACTAAAAACCA 601
        |||| || | | |||| ||| |||||| || ||| || |||||| ||||||||
Db 1052 AAAAAATAATAAATAATTTTAATCAATTTAACACAAAAAAAACTTATTATAATTA 993

Qy 602 CTTTAGAACAACGTGCT--ACTAACCTTGAAGGTTGTCATCAACTGCTTATAATCAAA 658
        ||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||||
Db 992 TTTCCAAAAAATAATAAAAAAAAAAAATAACTATAAACTCTATTAAAAATTAAAAA 933

Qy 659 TTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTTTAATAACTAAAACACTAG 718
        ||||| ||||| |||||||||| ||||||| ||| ||||||| ||||||| |||||
Db 932 TTAAAATAATTTATTATATAATCCTATAACCATCTAATTTTAAAAATACTACTATAAAACTA 873

Qy 719 ATCCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACACAGTTAATCGGAATA 778
        ||||| ||| |||||| || ||||| |||||||||||| |||||||||||||||||
Db 872 TTAACCTAACAATTACGACCTCTATATAAACTTTTAAAACTTTTTAAAAAATCCATAAA 813

Qy 779 TTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATTA 838
        |||| ||| ||||| || ||| ||| |||||||||||| ||| | ||||| ||||
Db 812 AAAAATATTTTTAATATAAAATAAAACTATAATACATAATCATAATAAAAACATTATTAAACCAAA 753

Qy 839 GTTTTATTAA 848
        |||| |||||
Db 752 AAATTATTAA 743

RESULT 12
US-10-239-676-102/c
; Sequence 102, Application US/10239676
; Publication No. US2003008269A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
;   FILE REFERENCE: 5013.1003
;   CURRENT APPLICATION NUMBER: US/10/239,676
;   CURRENT FILING DATE: 2002-09-24

Qy 302 ATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATC 361
        ||||||||| ||||||| |||| ||||||||| |||||| | ||| ||||||| |||
Db 1291 ACAAAACAAAATTAACAAAACTACTAAAAATCATCACACAATAAATCAAAAAAAAAATA 1232

Qy 362 TAATCAATGCTAAAGCGATGACATTAGCTTCACTACACAAGATCATGCCAAGATTGAAGCTA 421
        |||||||||| |||||||||||||||| | |||||||||| ||| | |||||||||||||
Db 1231 TTTAAAACCTTATATTTATATATTAATATACTAAATTTTAAAAATACTACTATAAAACTA 1172

Qy 422 GTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAG 481
        ||||||| | ||| |||||||| || || |||| |||||| |||| ||||||| ||
Db 1171 CAAAACAAAATAATTAATAAAAAATAATCAATTTAGAATCAGCCATCAACCAAGCTAATACGGATA 1112

Qy 482 AACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATA 541
        |||||||||||||| ||| ||| || ||||||||||| ||| |||||| ||||||||
Db 1111 TAAAACTTCAAAAATTTAAA-TAAAAAAACATAAAATTATTAATAAACATATACTTATA 1053

Qy 542 AAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGGACTAAAAACCA 601
        |||| || | | |||| ||| |||||| || ||| || |||||| ||||||||
Db 1052 AAAAAATAATAAATAATTTTAATCAATTTAACACAAAAAAAACTTATTATAATTA 993

Qy 602 CTTTAGAACAACGTGCT--ACTAACCTTGAAGGTTGTCATCAACTGCTTATAATCAAA 658
        ||||||||||| | ||| ||||||||||||||||||||||||||||||||||||||||
Db 992 TTTCCAAAAAATAATAAAAAAAAAAAATAACTATAAACTCTATTAAAAATTAAAAA 933

Qy 659 TTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTTTAATAACTAAAACACTAG 718
        ||||| ||||| |||||||||| ||||||| ||| ||||||| ||||||| |||||
Db 932 TTAAAATAATTTATTATATAATCCTATAACCATCTAATTTTAAAAATACTACTATAAAACTA 873

Qy 719 ATCCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACACAGTTAATCGGAATA 778
        ||||| ||| |||||| || ||||| |||||||||||| |||||||||||||||||
Db 872 TTAACCTAACAATTACGACCTCTATATAAACTTTTAAAACTTTTTAAAAAATCCATAAA 813

Qy 779 TTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATTA 838
        |||| ||| ||||| || ||| ||| |||||||||||| ||| | ||||| ||||
Db 812 AAAAATATTTTTAATATAAAATAAAACTATAATACATAATCATAATAAAAACATTATTAAACCAAA 753

Qy 839 GTTTTATTAA 848
        |||| |||||
Db 752 AAATTATTAA 743

RESULT 11
US-10-257-166-38/c
; Sequence 38, Application US/10257166
; Publication No. US2004023230A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
;   TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
;   FILE REFERENCE: 5013.1011
;   CURRENT APPLICATION NUMBER: US/10/257,166
;   CURRENT FILING DATE: 2002-10-07
;   PRIOR APPLICATION NUMBER: PCT/EP01/07470
;   DE 10032529.7
;   DE 10043826.1
;   PRIOR FILING DATE: 2001-06-29
;   2000-06-30
;   2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 38
;   LENGTH: 8136
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-38

Query Match   5.3%; Score 68.6; DB 17; Length 8136;
Best Local Similarity 44.1%; Pred. No. 0.016;
Matches 375; Conservative   0; Mismatches 471; Indels   4; Gaps   2;

```
;       PRIOR APPLICATION NUMBER: PCT/EP01/03968
;       DE 10019058.8
;       DE 10019173.8
;       DE 10032529.7
;       DE 10043826.1
;       PRIOR FILING DATE: 2001-04-06
;       2000-04-06
;       2000-04-07
;       2000-06-30
;       2000-09-01
;       NUMBER OF SEQ ID NOS: 228
;       SEQ ID NO 102
;       LENGTH: 11836
;       TYPE: DNA
;       ORGANISM: Artificial Sequence
;       FEATURE:
;       OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;       FEATURE:
;       NAME/KEY: unsure
;       LOCATION: (7603)
US-10-239-676-102

Query Match             5.2%;  Score 68.4;  DB 15;  Length 11836;
Best Local Similarity   44.3%;  Pred. No. 0.021;
Matches  371;  Conservative    0;  Mismatches  461;  Indels    6;  Gaps     2;

Qy        37 AAAAAACACAAAATCTTCTAACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAA 96
             || ||||  ||  | | |||  || |||||||| || || |||  | |  |||| |||||
Db      4357 ATAAAATATATATATATAAACTATATAAAAATATATATTATTATTATATATATATATTAT 4298

Qy        97 ATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTT 156
             || ||| | | | || |  ||| || ||||||||| |||| |  ||| ||  | |||||
Db      4297 ATAAATTTATATTTATATTATATAAATTTATATTTATAAATTTATTTATTATTTATTTATTA 4238

Qy       157 TTATTTTTAGTCATCTTTTAAGATATAAATATATCTTAATATTCTATGAATAAGAAAAGA 216
             || |||  | || | |  |||| ||| ||| ||   |||| ||  ||  |  ||| || |
Db      4237 TAAATTTTATATAAATATACTCTAATATATATTTTTACATTATATAATATATACTATATAT 4178

Qy       217 ATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCT 276
             || || |||||| |||| |  |  | ||| |  ||| || |  |||| || || | | ||
Db      4177 ATAAAAACAAACATATATATAAAAAATAATATACATATATAT--ATATATATATATATAT 4120

Qy       277 AGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAA 336
             || |  ||  | ||||| || ||| | | ||  || ||| ||| ||  |||  ||||||
Db      4119 ATATATAAAATCATTTATTCAACAATACCAAACTCTATCCTAATTACTT 4060

Qy       337 GCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTA 396
             || ||  ||  | || |||| | |  |||  ||  |||  | || ||  | || |||  |
Db      4059 ACCATACAACAATAAACAATATAAAAACCTCCCCCTCTAAATAAAAAAAAAAAAAAACAA 4000

Qy       397 CAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATGTGAAGCTGAAACAGTT 456
             |||| || || | || | || || ||| | | |  |  | ||| | ||| ||||| |||
Db      3999 TCAAATAA-----AATATAATATCATATAAATCAAACGCTATAAAAATAATAACCAAATA 3944

Qy       457 AACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAATCTAATTTAGAATCA 516
             ||| ||| |  | |||  || ||||||| ||| | |||  || | | ||||| ||||| |
Db      3943 AAAAAAAAAAATTTCATAAAAAAATTACTATTTATATTTTTAAAATTAATTTAAAAAAAA 3884

Qy       517 GCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTT 576
             ||| ||| | | |||  ||  || ||||  | | ||  || | || |||  ||||| || |
Db      3883 ACTTCGATATAAATAACATTTAAACAAAACTTACCGAAATAAAATAAACAACCATATAA 3824

Qy       577 GAAGCATAACAAAGCCTAAAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTTG 636
             |||| ||| |||| || |||  || ||||| ||||  |  ||| ||  || ||| || |
Db      3823 CTATATAAAAAAAAACTAATCTCAAAAATAACCATAAACAACAAAAACCTAAAAACCATA 3764

Qy       637 TCATCAACTGCTTATAAATCGCAATTCGATCATAAATGAATCTAAAAAAGCTAGTAAAGCTAGT 696
             ||| || || |||| | | || ||||| ||  ||| |  | || |||  ||| || |||||
Db      3763 AAATTCCTATCATATTAACAAAAAAAAAACCAATATACTAAAAACAAAATAAAATA 3704

Qy       697 AGTTTAATAACTAAAACACTAGATCTAAATGGGGAACGCCTTTAGATTCTAATGAG 756
             || |  |||  ||||| |||| || |||||  | ||| ||  | ||||  |||  ||
Db      3703 AAAAAAAATTATAAAAAATTAAAAATTAACAAAATTTAACATTCTTATAAAAA 3644
```

```
;       757 ATTACTACAGTTAATGGAATATTAATAATACGTTATCAACTATTATGAACAAAAGACT 816
             ||||| | |||| ||| || | ||| ||| || ||| |  ||| || |||| || || |
Db     3643 ATTATAAATTATACTTCTATATATAAAAACTCGACTATCAAAAATTAAAAAAATAAAATT 3584

Qy      817 AATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAA 874
             |||| | || | |||| ||||| ||||| ||| ||| |  ||||| | || ||||| ||
Db     3583 AAAAAACTATACAATAAAATTCATATAAAAATCCATTTATTAAACAAAAAAAAAAAAAA 3526
```

```
RESULT 13
US-10-240-453-114/c
;       Sequence 114, Application US/10240453
;       Publication No. US2003014832 6A1
;       GENERAL INFORMATION:
;       APPLICANT: OLEK, Alexander
;       APPLICANT: PIEPENBROCK, Christian
;       APPLICANT: BERLIN, Kurt
;       TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
;       TITLE OF INVENTION: Transcription
;       TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
;       TITLE OF INVENTION: With DNA Transcription
;       FILE REFERENCE: 5013.1009
;       CURRENT APPLICATION NUMBER: US/10/240,453
;       CURRENT FILING DATE: 2002-10-02
;       PRIOR APPLICATION NUMBER: PCT/EP01/03973
;       PRIOR FILING DATE: 2001-04-06
;       PRIOR APPLICATION NUMBER: DE 10019058.8
;       PRIOR FILING DATE: 2000-04-06
;       PRIOR APPLICATION NUMBER: DE 10019173.8
;       PRIOR FILING DATE: 2000-04-07
;       PRIOR APPLICATION NUMBER: DE 10032529.7
;       PRIOR FILING DATE: 2000-06-30
;       PRIOR APPLICATION NUMBER: DE 10043826.1
;       PRIOR FILING DATE: 2000-09-01
;       NUMBER OF SEQ ID NOS: 350
;       SEQ ID NO 114
;       LENGTH: 11836
;       TYPE: DNA
;       ORGANISM: Artificial Sequence
;       FEATURE:
;       OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;       FEATURE:
;       NAME/KEY: unsure
;       LOCATION: (7603)
US-10-240-453-114

Query Match             5.2%;  Score 68.4;  DB 15;  Length 11836;
Best Local Similarity   44.3%;  Pred. No. 0.021;
Matches  371;  Conservative    0;  Mismatches  461;  Indels    6;  Gaps     2;

Qy        37 AAAAAACAAAATCTTCTAACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAA 96
             ||||||  ||  | |  ||  || |||||||| || || |||  | |  |||| |||||
Db      4357 ATAAAATATATATATAAAATCCTAAATAAACTATAAAAATATTATTATATATATATTAT 4298

Qy        97 ATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTT 156
             || ||| | | | ||| |  ||| || ||||||||| |||| |  ||| ||  | |||||
Db      4297 ATAAATTTATATTTATATTTATATAAATTTATATTTATAAATTTATTTATTATTTATTA 4238

Qy       157 TTATTTTTAGTCATCTTTTAAGATATAAATATATCTTAATATTCTATGAATAAGAAAAGA 216
             || |||  | || | |  |||| ||| ||| ||   |||| ||  ||  |  ||| || |
Db      4237 TAAATTTTATATAAATATACTCTAATATATATTTTTACATTATATAATATATACTATATAT 4178

Qy       217 ATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCT 276
             || || |||||| |||| |  |  | ||| |  ||| || |  |||| || || | | ||
Db      4177 ATAAAAACAAACATATATATAAAAAATAATATACATATATAT--ATATATATATATATAT 4120

Qy       277 AGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAA 336
             || |  ||  | ||||| || ||| | | ||  || ||| ||| ||  |||  ||||||
Db      4119 ATATATAAAATCATTTATTCAACAATACCAAACTCTATCCTAATTACTT 4060

Qy       337 GCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTA 396
             || ||  ||  | || |||| | |  |||  ||  |||  | || ||  | || |||  |
Db      4059 ACCATACAACAATAAACAATATAAAAACCTCCCCCTCTAAATAAAAAAAAAAAAAAACAA 4000
```

```
Db   4059 ACCATACAACAATAAACAATATAAAAATAAAACCTCCCCTCTAAAATAAAAAAAAACAA 4000    Db   6685 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6626
                                                                                         
Qy    397 CAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTT 456    Qy    156 TTTATTTTTAGTCATCTTTAAGATATAAAATATCTCATGATAATTCATGATAAGAAGAAG 215
          |||| ||||||| |||||||||||||||||||||||||||||||||||||||||| ||           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3999 TCAAATAA---AATTAATAATATCATATATAACAAACGCTATAAAATAATAACCAAAATA 3944    Db   6625 AAAAAAAAAAAAAAAAAAAAAACGAAAAAAAAAAAATAAACATATATAAACCATATATA 6566
                                                                                         
Qy    457 AACAATAACCTTAAGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCA 516    Qy    216 AATCATCTTAAAGACATTAGTTGTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTC 275
          |||||||||||||||||||||||||||||||||||||||| |||||                    || ||| |||||||||||||||||||||||||||||||| ||||||
Db   3943 AAAAAAAAAAAAATTCTATAAAAAAAAAATTACTATTTTATTATATAAATTAAATAAAAAA 3884    Db   6565 AACAAACTAAAACCCTTATACACTACTATCCCAACATTATATAACTACAACCACTATAAA 6506
                                                                                         
Qy    517 GCCATCAACCAAGCTAATACGGATAAAAGACTTTTGATAATGAACACCCAAATTTAGTT 576    Qy    276 TAGGTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCAAACCCAATTAGA 335
          ||||||||||||||||||||||||||||||||||||| |||||||                    |||||||||||||||||||||||||||||||||||||||| |||||
Db   3883 ACTTCGATATATAAAAAACATTTAACAACAAAACTTACCGAAAAATAAAATACAACCATATAA 3824    Db   6505 AAACAATATAACAATTCCTCAAAAAATTAAACATAAAAATTACCACTATAAACAACAATC 6446
                                                                                         
Qy    577 GAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACCTTGAAGGTTTG 636    Qy    336 AGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGGCGATGACATTAGCTTCACT 395
          ||||||||||||||||||||||||||||||||||| |||||                    |||||||||||||||||||||||||||||||||||||||| ||||||
Db   3823 CTATATAAAAAAAAACTATAATCCAAAAATAATCCAAAAATAAACTAAAAAACTAATA 3764    Db   6445 ACACTTCTAAAATATATACCCAAAAAAACTAAAAACAAAAATTCAAACAAACATTTATACA 6386
                                                                                         
Qy    637 TCATCAACTGCTTATAATCAAATTCGCAATAATCGCAATAATTTGGATCTATACATAAAGCTAGT 696    Qy    396 ACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGT 455
          ||||||||||||||||||||| |||||||||||||||||||||||||||| |||||              |||||||||||||||||||||||||||||||||||||| |||||
Db   3763 AAATTCCTATCATCATATTAACAACAAAAAAAAACAATAATAAATCAATATAATAATAAAAATA 3704    Db   6385 CCAATACT---CAAAATAACATAAAATAGCTAAAAACTAAAAAAAACCACCCAAA 6330
                                                                                         
Qy    697 AGTTTAATAACTAACAACACTAGATCCACTAAAATGGGGAACGCTTTTAGATTCTAATGAG 756    Qy    456 TAACAATAACCTTAATGCAACATTAGAACAACATGCTAAAATGGCTAAAAACTAATTAGAATC 515
          ||||||||||||||||||||||||||||||||||||||||| |||||                    |||||||||||||||||||||||||||||||||| |||||
Db   3703 AAAAAAAAATTATAAAAAAAATTAAAATTAACAAATATATATATACAAAAAACAAAAAATAAATA 3644    Db   6329 TATTCATCAAC--AAATAAATATACAAAAATATAATACATACATAAATAATATATATTC 6272
                                                                                         
Qy    757 ATTACTACAGTTAATCGGAATATATTTAATAATACGTTATCAACTATTAATGAACAAAGACT 816    Qy    516 AGCCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGT 575
          ||||||||||||||||||||| |||||||||||||||||||||||||||| |||||              ||||||||||||||||||||||||| |||||||
Db   3643 ATTATAAATTATACTTCTATAATTAAAAAACTCGACTATCAAAAATAAAAAAAATAAAATT 3584    Db   6271 AACCATAAAAAAAAAATAAAAACTAAAATATATACT------ACAACATAAATAAAACT 6222
                                                                                         
Qy    817 AATGCTGATGCATTATCCAATAAAAATTCATATATAAGTTTATTAATAAAAAGTGATTCAAAATAATGAACAAA 874    Qy    576 TGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACACGTGCTACCTAACCTTGAAGGTTT 635
          |||||||||||||||||||| ||| ||||                    ||||||||||||| ||||||
Db   3583 AAAAAACTATACCAATAAAAATTCATATAAAAATTCATTTTATAAAACAAAAAAAAAAAAAA 3526    Db   6221 TTAAAACACTACTATCACAACCACTACAACAAATAACTCATTTTTATTATATATTTCC 6162
                                                                                         
                                                                                    Qy    636 GTCATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATCACAATAAGCTAG 695
                                                                                              |||||||||||||||||||||||||||||| |||||
                                                                                    Db   6161 ACTTATATAAAATACCAAAATAAACAAACTAATAAAAAATAAAAAAACAAAAATAAAATAA 6102
                                                                                         
                                                                                    Qy    696 TAGTTTAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCTTTTAGATTCTAATGA 755
                                                                                              ||||||||||||||||||||| |||||
                                                                                    Db   6101 AAAAAAAATTAAAAAAAAAAAAACAATTATAAAATATAAAAATTTCTATTTAAAATAATAATTA 6042
                                                                                         
                                                                                    Qy    756 GATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAGAC 815
                                                                                              ||||||||||||||||||||||||| |||||
                                                                                    Db   6041 AATACTAAAAAATAAATACTAATTAATAAATTACACAACTATAAATACACACCTAAAAATA 5982
                                                                                         
                                                                                    Qy    816 TAATGCTGATGCATTATCCAATAAAAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAA 874
                                                                                              |||||||||||||||||||| ||||
                                                                                    Db   5981 TCTAAATAATAAATTCCAAATATATATTTCACAATTAAAAAAAAAATAATAATTAA 5923
```

RESULT 15
US-10-221-613-101/c
; Sequence 101, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10011847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07

RESULT 14
US-10-311-455-1463/c
; Sequence 1463, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1463
; LENGTH: 8392
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1463

Query Match   5.2%;   Score 68.2; DB 15;   Length 8392;
Best Local Similarity   45.3%;   Pred. No. 0.019;
Matches 380;   Conservative   0; Mismatches 443; Indels 16; Gaps   3;

```
Qy     36 AAAAAAAACACAAATCTTCTAACAAAATCCTAAATAAAATCCTAAATAATAAATAAGCCGTTAAATTAACTAAAA 95
          ||||||||| ||||                               |||  |||||
Db   6745 AAAAAAAAAAAAAAAAAATGGTTTTCTTATCATCAACCAAAATTCTCTAGTAGTAATAAACGCTTATTATT 6686
                                                                          
Qy     96 AATTAAAAAAAATGGTTTTTCTTATCATCAACCAAAATTCTCTAGTAGTAATAAACGCTTATTATT 155
          ||| ||||                             ||  ||||
```

```
;    2000-06-30
;    2000-09-01
;  NUMBER OF SEQ ID NOS: 428
;  SEQ ID NO 101
;    LENGTH: 12405
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;  FEATURE:
;    OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;  FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (7895)
US-10-221-613-101
```

```
Qy          873 AA 874
                  |
Db          8468 TA 8467

Search completed: May  6, 2004, 00:19:34
Job time : 524.265 secs
```

```
Query Match      5.2%;   Score 68;   DB 13;   Length 12405;
Best Local Similarity   43.3%;   Pred. No. 0.025;
Matches 365;   Conservative  0;   Mismatches 475;   Indels  2;   Gaps  1;
```

```
Qy     35  TAAAAAAACACAAAATCTCTAACAAAATCCTAATAAATAAGCCGTTAAATTAACTAAA   94
               |  ||| || | ||| ||||||  ||  |   || ||| ||||||||||
Db   9308  TTAACATTCTAAATATATTTTAAAAATATAACAAACAACAAATCCCTAAAAAAAATCAA   9249

Qy     95  AAATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTAT   154
               |||||         || | ||      |||||| | ||  | |   |||||||
Db   9248  AAATAACTTCAATTTTTAATAAAAACAATTACAAATAAAATATACTACAAAAAAAACCAA   9189

Qy    155  TTTTATTTTTAGTCATCTTTTAAGATAAAATATATCT--TAATATTCTATGAATAAGAA   212
               ||| || |||||  ||   |   |  ||  ||  ||||| |||||  ||
Db   9188  TTTAAAACACTAAAATAATAAAAAAACTAACCTTCTAAACCTTCTAACATCCACAAAAAAAAA   9129

Qy    213  AAGAATCATCTTAAAGACTATTAGTTGTTAGGTACAACATCCTTTCTTAGCATTGGGAT   272
                 |  ||  ||||  | | ||   |  | ||| |||  ||    |||  || ||
Db   9128  CAAAACCACTAAACACGAAAATATATACTTAAAAAAAAATCCAAACTAAAAATATAAAT   9069

Qy    273  TTCTAGCTGTATGTCTATTACTAAAAAGAGATGCAAACCCAAATAATGGCCAAACCCAATT   332
               |  | || |    | ||      ||  | |    | |  ||| | ||
Db   9068  TCAATAATCTTCAACGAATAAACAATTTATACCACAAATAAAATCCAACTAATTTAACAT   9009

Qy    333  AGAAGCAGCGCGAATGGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTC   392
               |||||||||  || | ||         |||| || |  |   |||||
Db   9008  AAAAATAAACATAAAAAAATTAAAAACCAACTCAAAAAAATTTAAAATCCAAAA   8949

Qy    393  ACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAAC   452
               || || || |||| || |   |||||||| |   |||   |||||
Db   8948  AATAAAAACAACCTACCACTAAAAAATCAAAAAAAAATCAAAAATATATAAAATCTTAAAA   8889

Qy    453  AGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGA   512
               ||  ||||       | ||| |  |||     |||||| ||
Db   8888  TCCAAAAAAAAATAATATAAAAAATATAAAAATATTACTATATCAAAAACAAATT   8829

Qy    513  ATCAGCCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTT   572
               | || |||  ||||   |  ||     |  |||      || ||||
Db   8828  AATAAAAACTAATCAATAATTTACCTTTAAAAAAATCATAAAAATAACTTTAATCAAAAACTA   8769

Qy    573  AGTTGAAGCATACAAAGCACTAAAAACCCACTTTAGAACAACGTGCTACTAACCTTGAAGG   632
               | ||||  |   ||| | |  || |   |||||  |||   | ||||||
Db   8768  TTTTAACAAAAAATAAAACTAAAAATCTAAAAATCTTAAAAAAATAAAAAAATAAAAAA   8709

Qy    633  TTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTCATACAATAAAGC   692
               |||| || |      | | ||  ||||||| ||| |  |||
Db   8708  AAAACTAAAATAAACAACTCTTAAAAAATTCTACTACAAAAAATAAAAAATAAAAAC   8649

Qy    693  TAGTAGTTTAATAACTAAAACACTAGATCCACTCCACTAAATGGGGGAACGCTTTAGATTCTAA   752
               |||||||  | || |  |     |    ||||||||| ||| ||
Db   8648  AATATTTTTAAAAACTATAAAATCTCCAAAAATTTTTTTCCTTTAAAAATAAA   8589

Qy    753  TGAGATTACTACAGTAATCGGAATATTAATAACGTTATCAACTATTAATGAACAAAA   812
               |||| ||  |  || |  || ||     | || ||||||
Db   8588  AACTATTTCAACATATAAATATACTAATAAAATCTAATAAAAATAAAAAATCATT   8529

Qy    813  GACTAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACA   872
               |||||||  | | |||||  |||      ||  ||||||  | || |||
Db   8528  AACAATACAAAAAAAAAAAAAAATAACTACTAAAAATAACAATAAAATAAAAAAAACAAA   8469
```

```
          GenCore version 5.1.6
     Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     May  5, 2004, 10:50:12 ; Search time 2985.35 Seconds
                                (without alignments)
                      13063.796 Million cell updates/sec

Title:          US-09-901-572a-1
Perfect score:  1306
Sequence:       1 aaaaacatcagattgttaat........taaaatcgtttatcaggtt 1306

Scoring table:  IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:       27511289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters:    55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
     1:  em_estba:*
     2:  em_esthum:*
     3:  em_estin:*
     4:  em_estmu:*
     5:  em_estov:*
     6:  em_estpl:*
     7:  em_estro:*
     8:  em_htc:*
     9:  gb_est1:*
    10:  gb_est2:*
    11:  gb_htc:*
    12:  gb_est3:*
    13:  gb_est4:*
    14:  gb_est5:*
    15:  em_estfun:*
    16:  em_estom:*
    17:  em_gss_hum:*
    18:  em_gss_inv:*
    19:  em_gss_pln:*
    20:  em_gss_vrt:*
    21:  em_gss_fun:*
    22:  em_gss_mam:*
    23:  em_gss_mus:*
    24:  em_gss_pro:*
    25:  em_gss_rod:*
    26:  em_gss_phg:*
    27:  em_gss_vrl:*
    28:  gb_gss1:*
    29:  gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                    SUMMARIES

                  %
Result       Query
No.   Score  Match Length DB  ID        Description
----  -----  ----- ------ --  --        -----------
   1  79.4   6.1   1101   29  CNS0039G   AL063921 Drosophil
c  2  79.2   6.1   1200   13  BX437758   BX437758 Drosophil
c  3  78.2   6.0   1392   29  CG757503   CG757503 P052-4-C0
c  4  78.2   6.0   1531   29  CG748014   CG748014 P041-4-B0
```

```
c   5  78    6.0  1124  13  BX436282  BX436282 BX436282
    6  77.2  5.9  1101  29  CNS003BD  AL064091 Drosophil
    7  77.2  5.9  1200  13  BX437758  BX437758 Drosophil
    8  77.2  5.9  1201  13  BX446437  BX446437 BX446437
c   9  76.8  5.8  1348   9  CG749499  CG749499 P043-4-A0
   10  76.4  5.8  1201   9  AL565455  AL565455 AL565455
   11  76.2  5.8  1101  29  CNS00EVL  AL069706 Drosophil
c  12  75.8  5.8  1101  29  CNS000B8  AL063632 Drosophil
c  13  75.6  5.8   641  28  AQ946120  AQ946120 Sheared D
   14  75.4  5.8  1225  28  CC238324  CC238324 CH261-192
   15  75    5.7  1201   9  AL536104  AL536104 AL536104
c  16  74.8  5.7  1201  29  CNS0167M  AL106396 Drosophil
c  17  74.2  5.7   998  13  BX436885  BX436885 BX436885
   18  74.2  5.7  1201  13  BX439779  BX439779 BX439779
   19  73.4  5.6  1201  13  CNS021G2  AL176843 Tetraodon
c  20  73.2  5.6  1146  29  CG757970  CG757970 P053-2-C0
   21  73    5.6  1201  13  BX420717  BX420717 BX420717
c  22  71.8  5.5  1201  13  BX356896  BX356896 BX356896
c  23  71.4  5.5  1896  29  CG753083  CG753083 P048-1-C0
c  24  71.2  5.5   960  13  BX346155  BX346155 BX346155
c  25  71.2  5.5  1566  29  CG757757  CG757757 P053-1-D0
c  26  71    5.5   781  29  BX145762  BX145762 Danio rer
   27  71    5.4  1056  13  BX415058  BX415058 BX415058
   28  71    5.4  1091  13  BX424950  BX424950 BX424950
c  29  70.6  5.4  1056  13  BX415058  BX415058 BX415058
c  30  70.2  5.4  1491  29  CG753221  CG753221 P048-2-A0
c  31  70    5.4  1373  29  CG750869  CG750869 P045-2-E1
   32  69.6  5.3  1626  14  CF238805  CF238805 AGENCOURT
   33  69.4  5.3  1092  29  CNS020K7  AL175696 Tetraodon
   34  68.8  5.3  1101  29  CNS00EO7  AL069440 Drosophil
   35  68.8  5.3  1101  29  CNS016LI  AL106896 Drosophil
c  36  68.6  5.3  1074  28  BZ696936  BZ696936 SP_Ba009
   37  68.6  5.3  1201  13  BX343343  BX343343 BX343343
   38  68.6  5.3  1201  13  BX361152  BX361152 BX361152
c  39  68.6  5.3  1592  29  CG750135  CG750135 P04-3-D0
   40  68.4  5.2  1187   9  AL514085  AL514085 AL514085
c  41  68.4  5.2  1210  29  CG749728  CG749728 P044-1-C0
c  42  68.4  5.2  1353  29  CG744812  CG744812 P037-3-B0
   43  68.2  5.2  1098  12  BM803974  BM803974 AGENCOURT
   44  68.2  5.2  1200  13  BX414560  BX414560 BX414560
c  45  68.2  5.2  1528  29  CG753854  CG753854 P049-1-D0
```

```
                         ALIGNMENTS

RESULT 1
CNS0039G
LOCUS       CNS0039G    1101 bp    DNA    linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Oseogawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
```

LOCUS   BX437758   1200 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION   BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01
                5-PRIME, mRNA sequence.
ACCESSION   BX437758
VERSION   BX437758.1 GI:30773605
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1200)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
                Genoscope - Centre National de Sequencage
                BP 191 91006 EVRY cedex - France
                Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
                Library was constructed by Life Technologies, a division of
                Invitrogen. Contact : Peng Liang Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
                Faraday Avenue Genoscope Sequence ID : CSOCAP008CA01QP1.
FEATURES   Location/Qualifiers
     source   1..1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSOCAP008YB01"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

ORIGIN

Query Match   6.1%; Score 79.2; DB 13; Length 1200;
Best Local Similarity 30.8%; Pred. No. 0.0024;
Matches 242; Conservative 176; Mismatches 343; Indels 25; Gaps 3;

Qy   66 TAAATAAATAAGCCGTGAAATTAACTAAAAAATTAAAAAAATGGTTTTCTTATCAACCA 125
          ||| :|| || ||| || : | |||| ||||:::|||| : || : : |
Db   1043 TAATTWATHYTATTTWTCTATAWTAWATAWATAAAAHWWNTTTYTWAWAAAAAAAWWT 984

Qy   126 AAATTCTCTAGTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTAAGATATAAA 185
          | ||| || | |||||| ||| ||| | ||| ||||: |:| :|| || |
Db   983 TTTTWWYATATTAAAAAAWATTWTTWTCTTYTTCTTTTTYTWYWTWTWAAATW 924

Qy   186 TATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTGTTAGG 245
          |||| | || | ||| |||| | |||| | :|| :|| ||||| | ||| ||
Db   923 TTTTTATAWAAWTTTTATTTTTTTTCMWYYTYTTTATTTATTTTTYTMTTATA 864

Qy   246 TACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAGATGC 305
          || | | |||| |||| ||| ||||||||| | ||| |||| |||||| |||
Db   863 TATAAAWATATATTWYTYWTYTTTTTTTWTWAWATMTTYTCWW--------- 817

Qy   306 AAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCCGAATGGAGTAACAGATCTAAT 365
          |||||| | |: :|| | | |||:||| ::| ||| ||| : || ||||| |
Db   816 -----WAATTTTTTTTAWAAWWWTTTTMHATCTWTWAAAAAATAWWTATAAAAA 763

Qy   366 CAATGCTAAAGCGATGACCATTAGCTTCCACTACAGACTATGCCAAGATTGAAGCTAGTTT 425
          || ||| ||| | | |:|:||| | | | || ||: | | ||| | |:: ||
Db   762 HTCYTCYYYATAWWTWAAWATCYYTCMATAAAAAAAWTTYCTTCCYTTCTWWWTWT 703

Qy   426 ATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAAATGCAACATTAGAACA 485
          || | :| | ||| | || |::| | ||||||:||| || | | || |||||| |
Db   702 TTWAWANATYYMYTTCCWTWTTAWWAAWWTATMTAYYTCHYWWWAAACAWWAAAAAA 643

Qy   486 ACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACCGGATAAAAC 545
          | : || : || || || |||| | || || ||| | |:|| || |
Db   642 AAWAAAAAWAWAWMWHCYTTWAWAAAAAAATCCTYMWWATAAAMMAAAAWATMWT 583

RESULT 2
BX437758/c

RESULT 3
CG757503/c

```
LOCUS       CG757503              1392 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG757503
VERSION     CG757503.1  GI:37986131
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1392)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-3, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1392
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN
```

Query Match   6.0%; Score 78.2; DB 29; Length 1392;
Best Local Similarity   42.8%; Pred. No. 0.0032;
Matches 503; Conservative   0; Mismatches 658; Indels   13; Gaps   2;

```
Qy     36 AAAAAAACACAAAATCTTCTAACAAAATCCTAAATAAAATAAGCCCGTTAAATTAACTAAAA   95
          ||||||||       ||||||||    |||||||      |||||||      ||||||
Db   1392 ATAAAAAAATAAAATAAATAAAATAAAATAAATTAAAATATAAAATATAAAAATAATAAATA 1333

Qy   1056 TAGTTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTAGCAACTATGCTCATCAAC 1115
          ||||     ||||     ||||||||     |||||||    |||||||
Db    385 AAATAAAAATAAATAAAATAAAATAAAATAATAAAATATAAAAATAATAATAAATTAA  326

Qy   1116 TGGTTATTTATATTCCCTTATAAGTTGGTTAAAGCAGCTGATGCTAAACGTTGGATT 1175
          ||||||||      |||||||||
Db    325 TTTTATTTATTATTATATAAAAATTTATAAATTTATAATAAATAAATTTATATA  266

Db   1176 ACAATACAAATTAAATAAAATGGAAAATGTTCAACAA 1209
```

```
Qy     96 AATTAAAAAATGGTTTTCTTATCAACCAAAATTCTCGAGTAATAAAACGGCTTATTATTATT  155
          ||||||       |||       |||
Db   1332 CTTSTTTTTTWTWAAWAA-AAATATAWATAWTWWWTARAAAWAWWTAYCTCCCW  524

Qy    156 TTTATTTTAGTCAATCTTTAAGATAAAAAAAATATCTTATATCTTTATATTT  215
          ||||      ||||       |||
Db   1272 AANAATAATATTTAATAATAAAAAATATTAAAAAATAATTATTAAAATATTAT 1213

Qy    216 AATCATCTTAAAGACTATTAGTTTGTTAGTGTACACATCCTTCTTAGCATTGGGATTTC  275
          |||      |||||       ||||
Db   1212 AAAAAAAAAAAATATAA------AATAAAATAATAAAATAAAATAAAATAAAA 1165

Qy    276 TAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGA  335
          ::||       |||||||
Db   1164 ATATTATTATAAAAATAATATCTAAAAAATAAAAATAAAAATAAAA 1105

Qy    336 AGCAGCGCGAATGGAGTAACAGATCTAATCAATGCTAAAGCGATGCATTAGCTTCACT  395
          |||||       ||||
Db   1104 ATAAAAAAAATAATATAAAAAATAATAAAATAATAAAATAAAATATTAAAT 1045

Qy    396 ACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGT  455
          ||||       ||||
Db   1044 AAAATTATTAATATAAAAAATAAATAAAATAAAAAATAATAAAACTAATTTAGAAATC  515

Db    984 AAAATATAAAAAAAATATATAAAAAAAAAAAAATAAAAAAAATTAATATATATAATAATA  925

Qy    456 TAACAATAAAACCTTAATGCAACATTAGAACACTCAACTAAAAAAATGCTAAAAAT  515

Qy    516 AGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGT  575
          |||||
Db    924 TATAAAAAANAAAATAAAATTATAAAATTAAAATCAAAATATAATATAATATATATTAT  865

Qy    576 TGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTCGAAGGTTT  635
          |||||
Db    864 AAAAAAAAAAAAAAAATTAAAAATATATTAAAATAAATAAAATAAAATAATAATAATAA  805

Qy    636 GTCATCAACTGCTTATATCAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAG  695
          |||       ||||
Db    804 AAAATAAAAAAAAAATAAAAATATATAAAATATATTATTAATTAATATATATAATAAAATAAATA  745

Qy    696 TAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGA  755
          |||||
Db    744 AAAAAAAAAATAATAAAAATAAAAAATATAAAATAAAAAAAAAATATAAAAAAT  685

Qy    756 GATTACTACAGTTAATCGGAAATATTAATAATACGTTATCAACTATTAATGAACAAAGAC  815
          |||       |||||
Db    684 TATTAAATAAAATTATAATTTATTTTATTAATATAATTAAATATAAAATAATAAAA  625

Qy    816 TAATGCTGATGCATTATCTAATCATAAAAAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAAG  875
          ||       |||||
Db    624 ATATTTATAAAATAAAATAAAAAAATNTATAATAAAAAAAAAAAAAATATAAAAT  565

Qy    876 TTTTGTAGGGACTTTACAAACGCTAATGTCAATGTTCAAACCTTCAAACTACAGTTTGTTGCTT  935
          ||||
Db    564 TTTTAAATAAAATTTATATTTAAATTTAATTTTATTAATTATATATATATATTAT  505

Qy    936 TAGTGCTGATGTAACACCCGTCAATTATAAAAATGCAGAGGACCGTTGGAATGGTGA  995
          |||||
Db    504 ATAATAATATATAAAAAAAAAAAAAATAATTAATTAAAATATATAATATATAATTAAAAAAAATA  445

Qy    996 TGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTA 1055
          ||       ||||
Db    444 ATAATAAAAAAAAAAAAAATAAAT-AAATAAAATTAAAATAAAATAAAATATTATTTA  386
```

Db   849 AAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAANANANANAAAAAAAAAAN 908
         ||||| ||||
Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
             || |||
Db   909 AAAAAAAAAAAANNNAAAAAAAAANNNAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAN 968
         ||| |||||| ||| ||| |||
Qy   481 GAACAACTAAAAATGGCTAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACCGGAT 540
         ||| |||||| ||| ||| |||
Db   969 NAAAAAAAAAAAAAAAAAATAAAATAAATAAAAAAANNAAAAAAAAAAAAAAAAAAAAAA 1028
         |||||
Qy   541 AAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 600
         |||||
Db 1029 AAAAAAAAAAAAAANAAAAAAAAAAANANANAAAAAAAANNAAAANNAAAAAAAAAN   1088
         ||||
Qy   601 ACTTTAGACACACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATATCAAATT   660
         ||||
Db 1089 NAAANANANAAAAAANANNNAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1148
         |||| |||||
Qy   661 CGGAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
         |||| |||||
Db 1149 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAN 1208
         ||||| |||
Qy   721 CCACTAAATGGGGGAACCCTTTTAGATTCTAATGAGATTACACAGTTAATCGGAATATT   780
         ||||| |||
Db 1209 AAANNAAANAAAAAANAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1268
         |||| ||||
Qy   781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT   840
         |||| ||||
Db 1269 AAAAAAAAAAAAAAAAAAAAANAAANAAAAANNAAANNAAAAAAAAAAAAAAAAAAAAA 1328
         |||| |||
Qy   841 TTTATTAAAAAGTGATTCAAAATAATGAACAAA 874
         |||| |||
Db 1329 NAAANAAANAAAAANAAAAAAAAAAAAAAAAAAANA 1362

RESULT 5
BX436282/c
LOCUS   BX436282   1124 bp   mRNA   linear   BST 15-MAY-2003
DEFINITION BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YC01
            5-PRIME, mRNA sequence.
ACCESSION   BX436282
VERSION   BX436282.1   GI:30787521
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1124)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP001AB01QP1.
FEATURES   Location/Qualifiers
     source   1..1124
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP001YC01"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

Db   265 AATATATAAAAAAATATTTAAATAAATTTAAAATAA 232
         | ||||| ||| | ||||| ||| || ||| |

RESULT 4
CG748014
LOCUS   CG748014   1531 bp   DNA   linear   GSS 24-OCT-2003
DEFINITION P041-4-B02.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG748014
VERSION   CG748014.1   GI:37968940
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
ORGANISM   Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1531)
AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen.M. and Sommer,R.J.
TITLE   An integrated physical and genetic map of the nematode Pristionchus
            pacificus
JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE   22835951
PUBMED   12884007
COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..1531
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."

ORIGIN

Query Match   6.0%;   Score 78.2;   DB 29;   Length 1531;
Best Local Similarity   39.4%;   Pred. No. 0.0031;
Matches 344; Conservative   0; Mismatches 530; Indels   0; Gaps   0;

Qy   1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTCAACAA 60
         ||||| | ||| ||||
Db   489 AAAAANAAAAANAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAA 548
         ||||| | ||| ||||
Qy   61 AATCCTAAATAAAATAAGCCGGTTAAATTAACTAAAATTAAAAAAATGGTTTTTCTTATC 120
         |||| | |||||| ||||||
Db   549 AANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNA 608
         |||| | |||||| ||||||
Qy   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTAGTCATCTTTTAAGAT   180
         || ||||| || |
Db   609 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAANNAAAAAAAAA 668
         || ||||| || |
Qy   181 ATAAATATATCTTAATATTCTATGAATGAATAAGAAAAAGAATCATCTTAAAGACTATTAGTTTG 240
         | ||| ||| |
Db   669 AAAAAAAAAAAAAAAAAAAANAAAANNAAAAAAAANAAAAAAAAANNAAANAAAAAAAAAAAAAA 728
         | ||| ||| |
Qy   241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTGTATGTCTATTACTAAAAAA 300
         | | ||||| ||
Db   729 ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAANNAAAAAAANNAAAAAAAAA 788
         | | ||||| ||
Qy   301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCCGCGATGAGTTTACACAGAT 360
         | |||||||| ||||
Db   789 AAAAAAAAAAAAAAAAAAAAAAAAAANAAAAANRAAAAAAAAAAAAANARAAAAAAAAAAA 848
         | |||||||| ||||
Qy   361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGAATTGAAGCT 420

ORIGIN

ORGANISM   Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
           Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
           Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
  AUTHORS Genoscope.
  TITLE   Direct Submission
  JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
  COMMENT Determination of this BAC-end sequence was carried out as part of a
           collaboration with the Berkeley Drosophila Genome Project (BDGP).
           The BDGP is constructing a physical map of the Drosophila
           melanogaster genome using these BACs. For further information
           please see http://www.fruitfly.org The BDGP Drosophila
           melanogaster BAC library was prepared by Kazutoyo Osoegawa and
           Aaron Mammoser in Pieter de Jong's laboratory in the Department of
           Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
           NY. The library is named RPCI-98 and was constructed by partial
           EcoRI digestion of Drosophila DNA provided by the BDGP from the
           isogenic strain y2; cn bw sp, the same strain used for the BDGP's
           P1 and EST libraries. A more detailed description of the library
           and how to order individual BAC clones, the entire library, or
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K08"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

Query Match   5.9%; Score 77.2; DB 29; Length 1101;
Best Local Similarity   40.4%; Pred. No. 0.0052;
Matches 157; Conservative 47; Mismatches 185; Indels 0; Gaps 0;

Qy   13 TTGTTAATCTGATATCTTGCTTAAAAAACACAAAATCTCTAACAAAATCCTAAATAA 72
             |||| |:|| |:|| |||||| |||||||| ||||||
Db   664 TTATTTAWNAAAWATTAAAAAAAAAAAAAATATAAAAAATTATAAAAAWTW 723

Qy   73 ATAAGGCCGTTAAATTAACTAAAAAAATTAAAAAAATTAGGTTTTTCTTATCAACCAAAATTCT 132
             |||||||||| :||||||||| ||||||||||||:|||||||||
Db   724 TAAATATATAWTWTTTAWAAAATATATTAAAAAAAWATAWTTTTATTTTTATATTAAAAAATAT 783

Qy   133 CTAGTAATAAACGCTTATTTATTTTTTATTTTTTAGTCATCTTTTAAGATATATAAATATATCT 192
             :||||||| |||| |||:||| ::||:|||
Db   784 WATWATTTATATATNNWATAWTTTWTTTAAAATTWTTAAAWTWTTMATTAAAWATTN 843

Qy   193 TAATAATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACCAACA 252
             |||| |||||||||||||||||||||||||||| |||
Db   844 TAWAATTAAAAAWAAAAWAAATATAAAAAATAATAATAATWTATWTWTWAAATAAWTAAAW 903

Qy   253 TCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTTCTATTACTACTAAAAAAGATGCAAACCCA 312
             ||:|:|||||||||||||||||| |||||||||||||
Db   904 TAAWAAAAAAAAAATTAAWATAAAAATWAAWTAAWTTWTTTWMTYTAWAAAWTWATTAWAWAW 963

Qy   313 ATAATGGCCAAACCCAATTAGAAGCAGCAGCGCGGAATGGAGTTAACAGATCTAATCAATGCT 372
             |||| |:||| :: ||
Db   964 AATATTTTTATATWATWATTATATAATWAAAAATAAAATWAAAATNWAWAWWAYAA 1023

Qy   373 AAAGCGATGACATTAGCTTCACTACAGAGA 401
             |:|| :|:|
Db   1024 AWAYACWAAAAAWTWTATWYATAWMAAAA 1052

RESULT 7
BX437758
LOCUS   BX437758   Homo sapiens   1200 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01
            5-PRIME, mRNA sequence.

ORIGIN

Query Match   6.0%; Score 78; DB 13; Length 1124;
Best Local Similarity   24.2%; Pred. No. 0.0038;
Matches 203; Conservative 255; Mismatches 371; Indels 10; Gaps 2;

Qy   83 AAATTAACTAAAAATTAAAAAAATGGTTTTTTCTTATCAACCAAAATTCTCTAGTAATAA 142
             |:||:||| :|||||| |||||| || |:||||
Db   1122 AAYYYYYYAAAAAAAAAAAAAHHHYYYYYYHAAAAAAAHHAAAAAHYYHAYAAA 1063

Qy   143 ACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGATATATAAATATCTTAATATTCTA 202
             |:|::|: |:: .:|.: || |:||:|:
Db   1062 AYNYYYYYHAAAYYYAYYYYHAYYYYAAAAAAHYYYYYYYYYYYYYYHAAAAYYYYY 1003

Qy   203 TGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTA 262
             : || | |:   :|| ||| ||||:|:
Db   1002 YYYYYYAAAAAAAAAHMAAAYYYYMHYHHYYYYYYAAAAAAAAAYYYYYYHAA 943

Qy   263 GCATTGGGATTTCTAGCTGTATGTCTATTCTATTACTAAAAAGATGCAAACCCAAATAATGCC 322
             |:   :::|:   :::.|:| :.:::|
Db   942 AYYYYYYYYYYYAYYYYHAAAAHHHHHAAAAAHAAAHAAAAAHYYYY 883

Qy   323 AAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGA 382
             |:|| || ||
Db   882 YYYYHHHAAAAAAYYYYHHAAAAAYYYYYYYYYYYYYYAAAA 823

Qy   383 CATTAGCTTCACTACAAGACTACTACGATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTG 442
             : || .:::| :: || |::|
Db   822 AAAAAAYYYYYY------YYYYYVAAAAAAHAAAAAAYYYYAAAAYYYYYYYYYYAAAY 769

Qy   443 AAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAA 502
             ||:   .:::||| :|:
Db   768 YYYAAAMYHAAAAAAAAAAAAAHYYHAAAAAAAAAAAHYHAAAAAHYHAARAAAAAA 709

Qy   503 CTAATTAGAATCAGCCATCAACCAAGCTAATACGGATAAACGACTTTTGATAATGAAC 562
             |:::| ::   :: | ::|
Db   708 AAAHYYYYYYYYYYYYYYYYYYYHAAAAAAHHAAAAAAAAAA- 650

Qy   563 ACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACCACTTAGACACAAGTGCTACTA 622
             | ::   :: |::: | ||| :|: :|:
Db   649 ---AAAHYYYYYYYYYYHHAHHAAAAAAAAAAAAYYYYYHAHYYYYYYYYYAA 593

Qy   623 ACCTTGAGGGTTTGTCATCAACTGCTTATAATCAACTCGCAATAATTTAGTTGGATCTAT 682
             ||:||||||||||| |:::||| ||
Db   592 AAAHYHAMYYYYYYYYYHAAAAAAAAAAAAAHHYYYYYY 533

Qy   683 ACAATAAAGCTAGTAGTTTAATAACTAAAAACTAGATCCACTAGATGGGGGAACGCTTT 742
             :: |||||||| ||:|||
Db   532 HAAAAAAAAAAAAAAYYAAAHYYHAAHHAAAAYYYY 473

Qy   743 TAGATTCTAATGAGATTACTACAGTTAATCGGAATATTAATAACGTTATCAACTATTA 802
             | :|:|| || |||| ||:| ::
Db   472 YYYYAAAAAMAAAAAAYYYYYARAAAHHAAAAAAAAAAAAAAYYYYYYYYAAAAAHYYY 413

Qy   803 ATGAACAGAAAGACTAATGCTGATGCATTATCTAATAGTTTATTTAAAAAAGTGATTCAAA 862
             | ::| ||| | |: |: ::
Db   412 YYYYYYYYHAAAAAAAHHAAAAAAAHAAAHHHHYYYYYYHARAAAAAAAAAAAAH 353

Qy   863 ATATGAACAAAGTTTTGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTAC 921
             :   : | ::: | :   :::||
Db   352 HYYYHHYYYYYAAAHYYYHYYYYYYYYYYYHYCCCYYYYYCCCYYYYTWAAAAC 294

RESULT 6
CNS003BD
LOCUS   CNS003BD   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K08 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL064091
VERSION   AL064091.1 GI:4941847
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)

```
ACCESSION   BX437758
VERSION     BX437758.1  GI:30773605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CSOCAP008CA01QP1.
FEATURES             Location/Qualifiers
     source          1..1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSOCAP008YB01"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match     5.9%;  Score 77.2;  DB 13;  Length 1200;
Best Local Similarity   30.6%;  Pred. No. 0.005;
Matches  256;  Conservative 156;  Mismatches 418;  Indels    6;  Gaps    1;

Qy     16 TTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCAACAAAATCCTAAATAAAATA  75
          ||:·|··||||:::::·:||::|||·· ·||:|·|·||:::· ||·:|||::|
Db    365 TWTWTTTTATWWWWNDDDDAKCAWADAGAATWTTTTTTWWRAGRGARKRDGTTDTRAADWW  424

Qy     76 AGCCGTWAATTAACTAAAAAAATAAAAAAATGTTTTCTTATCAACCAAAATTCTCTA  135
          ·||:::::· ||·| ||··||::||:||||:| ||||||||||||·||·|||||
Db    425 WWTWAAWAWWWDKRKTTTTTTTTKKRRAAAATTTTTTTTTTTWTWAKAAAARAAATW  484

Qy    136 GTAATAAACGCTTAATTTATTTATTTTTATTTTAGTCATCTTTAAGATAGTAAATATATCTTAA  195
          |·||::·:· :·||·|·|·||||· |·|||::||||||||||:|·|:||||
Db    485 TTWWRAAAAAAWWTWWWAARGRARGATTTTTTTAAWWWGGAGRTAWWTWTWTTTTTT  544

Qy    196 TATTCTATGAATAAGAAAGAAATCATCTTAAAGACTATTAGTTGTTAGGTACAACATCC  255
          ||··|·|· :|·|| :||:||::||||||·|:||·|||
Db    545 TTTTTTTTAWWWATWATATTTTTWTWAWAAASAGAWKATWTTTKKTTTA------  598

Qy    256 TTTCTTAGCATTGGGATTTCTAGCTGTATGCTATTACTAAAAAGATGCAAACCCAAAT  315
          ||::· :||·|||||||||||||·|·|||||·|·|||
Db    599 TWWKRAGGATWTTTTTTTTWTWAARGDWKWTWTWTTTTTTTTTTTWTTGTTTWWW  658

Qy    316 AATGGCCAAACCCAATTAGAGCAGCGGCGGATGAAGTTAACAGATCTAATCAATGCTAA  375
          ·||·||· :·||·:::· ·||·||||||:|·|:: ·||
Db    659 RDGARRTAKATAKWWTWTWTAWAWAGGAARKRRATWTWTWAARAWAWWAGARAAARGAGA  718

Qy    376 GCGATGACAATAGCTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCT  435
          ||||||· ||·|·||·|·||||:· :·:|||
Db    719 AGRAAWTTTTTTATTKAGARRGATWTTAWWWTATRARGARGADTTTTTATAWWTATTT  778

Qy    436 TATAGTGAAGCTGAAACAGTTAACAATAACCTTTAATGCAACATTAGAACAACTAAAAATG  495
          |·|||||·||·::·|| :|||::· :· ||||·|
Db    779 TTTWAWAAGATDKAAAAAWWWWTWTWAAAAATWWWAGARAAAKATWTTWAWAWAAA  838

Qy    496 GCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGAT  555
          ||||:· :||· :|||||· ||||·|||| :|·|·
Db    839 AAAAAWRARWAAATATATWTTATATAAKARAAAAAATAAAATAAAARARARRWKGAAAA  898
```

```
RESULT 8
BX446437
LOCUS       BX446437               1201 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  BX446437 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA006ZG10
            3-PRIME, mRNA sequence.
ACCESSION   BX446437
VERSION     BX446437.1  GI:31025727
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 7563.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CL0BA006ZG10FP1&cluster=7563.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CL0BA006ZG10FP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CL0BA006ZG10"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match     5.9%;  Score 77.2;  DB 13;  Length 1201;
Best Local Similarity   35.9%;  Pred. No. 0.0049;
Matches  265;  Conservative  81;  Mismatches 393;  Indels    0;  Gaps    0;

Qy    147 TTATTTATTTTTATTTTTAGTCATCTTTAAGATATAAAATATATCTTAATATTCTATGAA  206
          |··|||||||·|||||·||:||··| |:||:: ||| |·||
Db    101 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTWWWTWTWTTTWWWTTTWAWWAAWWAAA  160

Qy    556 AATGAACACCCAAATTAGTTGAAGCATACAACAAGCACTAAAAACCACTTTAGAACACACGT  615
          ·|||:|:|· ||:|··||:|||||||·|||·||| ||::||||::||||
Db    899 AAAAATAAAAWWTTTWTATAAAAAAWATTTWWAWWRAAARRAWAWWAAAAAAGGARAGAW  958

Qy    616 GCTACTAACCTTGAAGGTTGTGTCATCAACTGCTTATATACTTAYAYATCGCAAYAATTAGTG  675
          ||·|· |||:·:··||· · ::· |·
Db    959 AAWAATWTTTTTAATATRWAAAAAWTTTTTTTWTARAAAAWWDTTTATWTWTATWT  1018

Qy    676 GATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAAATGGGGGA  735
          |·|·:·|:||:|::||||||·| |:::·| |·
Db   1019 AWTATAGWAAATARDATWRAATTATWADAARRAAAAGAWATAKDAGATAAAAWTTTTTT  1078

Qy    736 ACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATTAATAAATACGTTATCA  795
          ||||::· ::||:::· :·|· ·|
Db   1079 TTTTTTTTWTAAAAAGAAAATWTTARAAAAAAAAAAWATAAAAAAAAAAAAAAAA  1138

Qy    796 ACTATTAATGAACAAAAGACTAATGCTGATGCATTACTAATAGTTTTATTAAAAA  851
          |·||· ·||·|:|||· ·|·
Db   1139 AAGAAKADDAAWAAGAAAGRAARAWAAAAAAWTAARA  1194
```

Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES   Location/Qualifiers
    source   1..1348
                  /organism="Pristionchus pacificus"
                  /mol_type="genomic DNA"
                  /strain="California"
                  /db_xref="taxon:54126"
                  /clone_lib="Ppa EcoRI BAC Library"
                  /note="The library was generated by a partial digest of
                  the genomic DNA with EcoRI and cloning into the BAC
                  vector."

ORIGIN

Query Match   5.9%; Score 76.8; DB 29; Length 1348;
Best Local Similarity 44.7%; Pred. No. 0.0054;
Matches 376; Conservative 0; Mismatches 457; Indels 8; Gaps 2;

Qy   34 TTAAAAAACACAAAATCTCTAACAAATCCTAAATAATAAGCCGTTAAATTAACTAA   93
           |||||| | ||| | || | |||||||| |||||
Db   1254 TTATAATTATTATTTTATAAATAAAAAAAAATATAAAAATTAATTAATATAAAATAT 1195

Qy   94 AAAATTAAAAAAAATGGTTTTCTTATCAACCAAAATTCTCAGTAATAAACGCTTATTA 153
           ||| || ||||||||
Db   1194 ATAAAATAAAAANAAAAAAATTATTNTATAAATATATTTAAATAATATATTTTA 1135

Qy   154 TTTTTATTTTAGTCATCTTTAAGATATAAATTTAGTGGATCTATACAA 213
           |||| | ||| | |||||
Db   1134 TATATAATAATAAAAATAATATATTAAAAAATATATTAAAAAAAATTAA 1075

Qy   214 AGAATCATCTTAAAGACTATTAGTTTGTTGTAGGTACACATCCTTCTTAGCATTGGGATT 273
           ||||| | ||| |||||
Db   1074 TAAATAAATAAAAAAAATATAAATTAATTATTTATAAAAAATAAAAATAAAAAA 1015

Qy   274 TCTAGCTGTATGTCTATTACTAAAAAGAGATGCAAACCCAAATAATGGCCAAACCAATTA 333
           | || | |||||
Db   1014 TTAAAATAAATAATTTAAAAATTAATAAAAATTATTAAAAAATTATTAATTAAT 955

Qy   334 GAAGCAGCCGCGAATGGAGTTAACAGATCTAATCAATGCGGATGCATTAGCTTCA 393
           ||| |
Db   954 TAAAAAAATAATAATAAAAAATAATAAAAATAATTATATAAAAAAAAAAAAAAAAA 895

Qy   394 CTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACA 453
           | |||
Db   894 AATAAAAAAATATAAANAATATAAAAATATTATATAAAATAATAA---AAAA 839

Qy   454 GTTAACAATAACCTTAACTGAACATTAGAACAACTAAATAAAATGGCTAAAAAACTAATTTAGAA 513
           ||||||||||||||||
Db   838 AAAAAAAAAAAATTTAAATAAAATTAAATAAAATATAAAAAAATAAATAAAAATATAAAAATA 779

Qy   514 TCAGCCATCAACCAAGCTAATACGATAAAACGACTTTTGATAATGAACACCCAAATTTA 573
           ||||||
Db   778 AAAAAAATAATAAATTTATTATAATATATTTTTAAAAAATTAAAANAAAAATAATAAA 719

Qy   574 GTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACACGTGCTACTAACCTTGAAGGT 633
           |||||
Db   718 AAAAAAAAAAATATATATATAATATATAAAAATTATAAAAAAATAAAAAATAAAAA--- 662

Qy   634 TTGTCATCAACTGCTTATAAATCAAATTCGCAATAAATTAGTGGATCTCACAATAAAAGCT 693
           ||||||
Db   661 -TAATAATAAATTAAAAATAAATAATATAAAATTATAAAATTATATTATAAAAAAAAT 603

Qy   694 AGTAGTTTAATAACTAAACTAAAACTAGTACCCACTAAATGGGGGAACCGCTTTTAGATTCTAAT 753
           |||||
Db   602 ATAAATAAAAAAAAATAAAAAAAAAAATATAATAATATNTAAAATAAAAA 543

Qy   754 GAGATTACTACAGTTAAATCGGAATATTAATTAATACGTTATTAATAATGAACAAAAG 813
           ||||
Db   542 ATAAATAATAAATAAATAATATAAAATAATAATAATATAAAAAAAAAA 483

Qy   814 ACTAATGCTGATGCATTATCTCAATAAGTTTTATTAAAAGTGATTCAAAATAAGAACAA 873

Qy   207 TAAGAAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACACATCCTTCTTAGCAT 266
           ||||| :: :| :|| :: ::: :|
Db   161 AAAAAAAAAAAAWAWAWWAAAAAAAAWAAAAWAWWTWWWWTWAAAAAA 220

Qy   267 TGGGATTTCTAGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAAC 326
           ||| :::| |||
Db   221 AAAAAAAAWAAAAWTTTAAAAWAAAAATAAAAAAAATAAAAAAAAAAAAAA 280

Qy   327 CCAATTAGAAGCAGCCGCGAATGGAGTTAACAGATCTAAAGGCTAAAGCGATGACATT 386
           ||:|
Db   281 AAACCCAAAAAAAAAAAAAAAACCMAAAAAAACMCCCCMAAAAAAACC 340

Qy   387 AGCTTCACTACAAGACTATGCCAAGACTAGTTTATCATCTGCTTATAGTGAAGC 446
           |
Db   341 AACACAAAAAAAAAATATAAAAAAAAAAAATWAAAAWAAWAAAA 400

Qy   447 TGAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAATGGCTAAAACTAA 506
           ||| :
Db   401 AACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATT 460

Qy   507 TTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGAATAATGAACACCC 566
           ||| |||
Db   461 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAA 520

Qy   567 AAATTTAGTTGAAGCATACAAAGCACTTAGAACAACGTGCTACTAACCT 626
           | :: ||| :||| | :| :::|
Db   521 AAAAAAAAAAAAAHAAAAACYTTTTTTTCCCHAAWTTYTTTTTWA 580

Qy   627 TGAAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAA 686
           | :||
Db   581 WWWTWTTTBTAAAAWRAATCYTMAAMCCCYCYTWAAAAAACYYCAMHYMAAAAAAA 640

Qy   687 TAAAGCTAGTAGTTTAATAACTAAACACTAGATCCACTAAAAGTGGGGAACGCCTTTTAGA 746
           ||| : |||| : :||
Db   641 AAAAAWCCCCYTTTTYYTTTYTTTTMAVAMCTTANCBCCHWWWAMCYYTTSBACCCCCCWCYT 700

Qy   747 TTCTAATGAGATTACTACAGTTAATCGGAATAATTAATACGTTATCAACTATTAATGA 806
           || | ||||
Db   701 TTTYCWCTYCWTTTYCCYTAAAKYCYMRAAWWAMAAWMAMACTCTWWAAAAMCTCTW 760

Qy   807 ACAAAAGACTAATGCTGATGCATTATCTCAATAAGTTTTATTAAAAAGTGATTCAAAATAA 866
           |||||||
Db   761 AAAAAAAAAAAANNNAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 820

Qy   867 TGAACAAAGTTTTGTAGGG 885
           || ||| : ||:|| :||
Db   821 AAAAAAAAAARADWTTTSTKG 839

RESULT 9
CG749499/c
LOCUS   CG749499   1348 bp   DNA   linear   GSS 24-OCT-2003
DEFINITION P043-4-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG749499
VERSION   CG749499.1 GI:37970425
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1348)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE   An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany

```
Db   482 TAAAATAAAAAAAAATTAAAAAAAAAAAAATAAAATAATTAAAAAAAAAATTAAAAAA 423
             |||            |||              ||||||  || ||| |  ||||
Qy   874 A 874

Db   422 A 422
```

```
Qy   297 AAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGGCGCGAATGGAGTTAAC 356
             |:   |:   : :   ||  :||                   
Db   670 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 729

Qy   357 AGATCTAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGA 416
             |: |::|      |::                     
Db   730 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 789

Qy   417 AGCTAGTTTATCATCTGCTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAAC 476
             |:|                       
Db   790 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 849

Qy   477 ATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATAC 536
             |||:|  |:   : ||:||                   
Db   850 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 909

Qy   537 GGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAA 596
             |||:|                       
Db   910 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 969

Qy   597 AACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCA 656
             |:|:|  |                      
Db   970 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 1029

Qy   657 AATTCGCAATAATTTAGTGGATCCATACAATAAAGCTAGTAGTTTAATAACTAAAACACT 716
             |::||                        
Db  1030 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 1089

Qy   717 AGATCCACTAAATGGGGAACGCGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAA 776
             |: |: : |                      
Db  1090 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 1149

Qy   777 TATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 818
             |||  | :  |||   :|| :| :||||||:          
Db  1150 ATAWATAWATAWATAWATAWATAWATATAWATAWATAWATAW 1191
```

RESULT 10
AL565455

```
LOCUS       AL565455    1201 bp    mRNA    linear    EST 12-MAY-2003
DEFINITION  AL565455 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF005YO18 3-PRIME, mRNA sequence.
ACCESSION   AL565455
VERSION     AL565455.2  GI:30549492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12916848.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9232.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF005BH09NP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF005YO18"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
```

```
Query Match          5.8%; Score 76.4; DB 9; Length 1201;
Best Local Similarity  31.8%; Pred. No. 0.0066;
Matches  242; Conservative 122; Mismatches 398; Indels 0; Gaps 0;

Qy    57 ACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAATCGTTTTTCT 116
             |:|:       |:|:||:|:            :|:|| |:|:     |:|:|:      
Db   430 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 489

Qy   117 TATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTA 176
             |||  |                        
Db   490 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAWATAW 549

Qy   177 AGATATAAATATATTCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAGACTATTAG 236
             |:|||:|:|:|:|:                      
Db   550 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 609

Qy   237 TTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAA 296
             |:|:                        
Db   610 ATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATATAWATAW 669
```

RESULT 11
CNS00EVL

```
LOCUS       CNS00EVL    1101 bp    DNA    linear    GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
```

```
source          1..1101
                /organism="Drosophila melanogaster"
                /mol_type="genomic DNA"
                /db_xref="taxon:7227"
                /clone="BACR29923"
                /clone_lib="RPCI-98"
                /note="end : T7"
ORIGIN

Query Match      5.8%;   Score 76.2;   DB 29;   Length 1101;
Best Local Similarity  32.0%;   Pred. No. 0.0074;
Matches  209;   Conservative 121;   Mismatches 321;   Indels  2;   Gaps   1;

Qy   42 ACCAAAAATCTCTRACACARAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAA 101
          ::::::|     :: ::  ::::|||:   :|   :|||:||| :|||: |
Db  443 MMMWMHAMATTYTCTCAHTWTMNMMWWAAITWTWTWNAAAWAAAWTTATWAATWAAAAAWA 502

Qy  102 AAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATT 161
          ::::|| |||:: :||:    :::|| |||  |: || |:  ||||    ||
Db  503 WWWWATTTTTWWWWTWTTTTWTAWWTWTTWAWWAAAAAAAAAWATAATTTAAAWWAATA 562

Qy  162 TTTAGTCATCTTTTAAGATATAAATAAATATCTTAATATTCTATGAATGAAGAAGAATCAT 221
          :  |:|| :|||      ||||  :||| |::||:|:|||||||||||||||||||||
Db  563 WATTAAWAATTTAWAAWWTATATTAATWTAATAWTWWATTATAATAAAAAAATATTTTT 622

Qy  222 CTTAAAGACTATTAGTTTGTTAGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTG 281
          :||||||      :||     : ||:   ||  |:  :|||: :|| |||
Db  623 TWATAAAATTTTTAATAATTTAATTAWTTATTAAATAAWTWATTWWWTAAAATAAATT 682

Qy  282 TATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGC 341
          :  :|| ||:|||||||||||||||||||:|||:   ::::||::||:|||||||||||
Db  683 TWAAATAWAAAAAAAAAAAAAAAAWATWAAAAAATWATAWATAAWTWATAAAAWAATAA--AAW 740

Qy  342 GCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGA 401
          |||||||     ||: |::  ::: ::|:|||| : :|| |   ||| ::|| :
Db  741 AAWAATWAWWWATAATATWWATATATTWTAWWWATWTWAWWWATWAWATAWWAAAW 800

Qy  402 CTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAGCAGTTAACAA 461
          :|: |:|   || || ::|::|:| ||  :| | |:::|||:: ::|:||||:||||||
Db  801 AWWATAWATWAWTAWAAAWAWAAAWAWAWATATAWWATWAATATATATTWAA 860

Qy  462 TAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGAATCAGCCAT 521
          :||::|:|||   ::||:||  |||||  |: ::|||::||  ||||||||:|||||:|
Db  861 TATWAATWATAAWAAAAAATAAWWTTTTWWAWWTATATAAAWATAWAAAAWAAA 920

Qy  522 CAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGC 581
          |||||||||||     ::::|  :||||:| ::|::|::    :|:||||||:||:|||
Db  921 AAAAAAAAAAATAAWANWWWWTWATATWTWTTATTAAAWWTWATATTWATTWWWATWWTATAW 980

Qy  582 ATACAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATC 641
          ::|: |::||  |      | |::|||:|| :|   |||  |:   ||||||::
Db  981 TTWTATWTATATWTWTAWTAWATATTTATWAWWTATATTTTAWAAAWTAATATATW 1040

Qy  642 AACTGCTTATAATCAAAATTCGCAATAAAATTAGTGGATCCTATACAATAAAGCTA 694
          |||||||||   ::|||  ||:| | | :|:||:|  ||||:  :|: |||||:|
Db 1041 ATAWWWTAWATATWWAWWAATTAWWTTATTATATATWTAAWATAWWAAAAAAAWWAAAW 1093
```

```
AUTHORS   Genoscope.
TITLE     Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
          collaboration with the Berkeley Drosophila Genome Project (BDGP).
          The BDGP is constructing a physical map of the Drosophila
          melanogaster genome using these BACs. For further information
          please see http://www.fruitfly.org The BDGP Drosophila
          melanogaster BAC library was prepared by Kazutoyo Osoegawa and
          Aaron Mammoser in Pieter de Jong's laboratory in the Department of
          Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
          NY. The library is named RPCI-98 and was constructed by partial
          EcoRI digestion of Drosophila DNA provided by the BDGP from the
          isogenic strain y2; cn bw sp, the same strain used for the BDGP's
          P1 and EST libraries. A more detailed description of the library
          and how to order individual BAC clones, the entire library, or
          filters for hybridization from the BACPAC Resource Center can be
          found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES            Location/Qualifiers
  source      1..1101
              /organism="Drosophila melanogaster"
              /mol_type="genomic DNA"
              /db_xref="taxon:7227"
              /clone="BACR01A24"
              /clone_lib="RPCI-98"
              /note="end : TET3"
ORIGIN

Query Match      5.8%;   Score 75.8;   DB 29;   Pred. No. 0.0086;   Length 1101;
Best Local Similarity  38.4%;
Matches  198;   Conservative  67;   Mismatches 249;   Indels  1;   Gaps   1;

Qy   59 AAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAATGGTTTTTCTTA 118
          ||||     ||: |        :::|||||||||||||||||||||||||: |
Db 1092 AAAAAAWAAAAATTAAATATTTWWAAAAAAAAAATTATAAAWATWAATWATWWAWTA 1033

Qy  119 TCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTATTTTTAGTCATCTTTAAG 178
          :|: |:|:||     ||| |:| ||  |||||  || ::|::|:
Db 1032 ATAAAAAAATAAAATTWTAATATAAAAATWTAAWATATAAAATATWTTAATAAATTAAAW 973

Qy  179 ATATAAAATATATCTTAATATTCTATGAATAGAAGAAGAATCATCTTAAAGACTATTAGTT 238
          ||| :|:|: |  :|||   :::| |: |||| |:: || |::||
Db  972 ATAWATWAWAAAAAATAAAAAAAAWWAAAAATAAATAAWTWTWWAAWTTWTAAWWATWAAAAWAT 913

Qy  239 TGTTAGTACAACATCCTTCTTAGCATTGGGATTGTCTATTACTAAAA 298
          ||| ::|||||||||||||||||||||||||||:|: ||:| ::|||
Db  912 ATAATTTAAAWWATWAAAAWTTAAWWTWWAAWTTTWTTTTTTTWTTAAWTTAAAW 853

Qy  299 AAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAGAGCTTCACTACACAGACTTAACAG 358
          |:: ||:|||||:|: |:   ::|:   ::|:|:|| |:|||||||||||
Db  852 AAAAAWTAAAAAAATAAAATWAAAAAWWWAAAAWWWAAAWWATAAWWATAWWWAAAAAAAA 793

Qy  359 ATCTA-ATCAATGCTAAAGCGGATGACATTAGCTTCACTACACAGAGATGCCAAGATTGAA 417
          :||:  |: :|| | | :|::|:||||||||:::| ::::::||||| |||||::|
Db  792 HTAAATWAAAATTACMAAAATYAAATAAAAWTWATTTWWAWAWWATWTWAWAWWWATAWAW 733

Qy  418 GCTAGTTTATCATCTGCTTATAGTAGTGAGCTGAAGCTGAAGAGTTAACAATAACCTTAATGCAACA 477
          |||:||  ::|::| |||||| ||| || :|::||::|||||||| |::
Db  732 AAAAAATTAWATTHCMAAATWAAWCAAAAHAAWAMCAMAAAAAAATATTTTYCCAAWA 673

Qy  478 TTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCGCCATCAACCAAGCTAATACG 537
          :|:::: ||:|||| || ||||||:|| || :||: :|::|
Db  672 TCACAAYMACCAACACAAAAATCAMACHAAACACCAAYYAAWCAMAAYCACAAATYBCAAAT 613

Qy  538 GATAAAAACGACTTTGATAATGAACACCCAAATTT 572
          ||||||| :||  ||:|||||| |:|:|||||:||
Db  612 CAYYAAAGRSDCYYAAAYAAAYKSYHHAWATTTT 578
```

```
LOCUS       AQ946120                 641 bp    DNA     linear   GSS 27-JAN-2000
DEFINITION  Sheared DNA-46J23.TR Sheared DNA Trypanosoma brucei genomic clone
            Sheared DNA-46J23, genomic survey sequence.
ACCESSION   AQ946120
VERSION     AQ946120.1  GI:6769385
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,B., Malek,J., Fujii,C.,
            Gerrard,C., Leech,V., de Jong,P., Uilu,E., Melville,S.,
            Donelson,J., Fraser,C. and Adams,M.
  TITLE     Determination of clone end sequences from Trypanosoma brucei GUTat
            10.1 sheared DNA library
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: Sheared DNA-46J23.TF
            Contact: Najib M. El-Sayed
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: nelsayed@tigr.org
            Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
            DNA library constructed at TIGR. Clones will be available for
            distribution through ATCC. Sheared DNA end sequences search page:
            http://www.tigr.org/tdb/mdb/tbdb/.
            Seq primer: M13-Reverse
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..641
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927/4 GUTat 10.1"
                     /db_xref="taxon:5691"
                     /clone="Sheared DNA-46J23"
                     /clone_lib="Sheared DNA"
                     /note="Vector: pUC18; Site 1: SmaI; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from a cloned population of
                     Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                     sheared to give a tight size distribution (approx 2 kb).
                     The v + i method used for the library construction is
                     described in detail in Smith, H.O. and Venter, J.C.
                     (Making small insert libraries for whole genome shotgun
                     sequencing projects. In Genome Sequencing: A Practical
                     Approach, eds. M. Vaudin and B. Barell, Oxford University
                     Press, 1999)."
ORIGIN

Query Match             5.8%;  Score 75.6;  DB 28;  Length 641;
  Best Local Similarity 46.0%;  Pred. No. 0.012;
  Matches  255;  Conservative    0;  Mismatches  299;  Indels    0;  Gaps    0;

Qy        750 TAATGAGATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACATTAATGAACA 809
              ||||  || ||||||| ||   |         |||||||||    || |||||||
Db        624 TAGTAATAGTTACTACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 565

Qy        810 AAAGACTAAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGA 869
                              |||     ||||| ||| |||     ||||||||||||||
Db        564 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 505

Qy        870 ACAAAGTTTTGTAGGGACTTTTACAAGCGCTAATGTTCAACCTTCAAACCTACAGTTTTGT 929
                             |||||     |||  |||||||     |||||| ||||||||||||
Db        504 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 445

Qy        930 TGCTTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAA 989
                ||||||  ||||     |||  |||||||||||||||||||||||||
Db        444 TAATAATAGTTACTACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 385

Qy        990 TGGTGATGAACCTCAAGTAGAATTCTTGCAAACACGAAGTATCACAGATGTTTCTTG 1049
```

```
Db        384 TAATATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA

Qy       1050 GATTTATAGTTTAGCTGGAACAAACGAGCTACCAATTTAGTTTTAGCAACTATGGTCC 1109
                |||||     |||  |||       |||  ||||||||||||||||||||||||||
Db        324 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 265

Qy       1110 ATCAACTGGTTATTTATATTCCCTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGT 1169
                 |||  |||||||||||||     |||   |||||||||||||||||||||||||
Db        264 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 205

Qy       1170 TGGATTACAATACAAATTAAATAATGAAATGTTCAACAAGTGAGTTGGCCACTTCAAC 1229
                   |||||||||||||||||     |||||     ||||||    |||||||||||
Db        204 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 145

Qy       1230 TAGTGCAAATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAA 1289
                 |||      |||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TAATAATAATAATAATAATAATAATAATAATAATGATGATGATAATAATAATAATAG 85

Qy       1290 AATCGTTTATCAG 1303
                |||  |||||||
Db         84 AAGAGTGTTTGAG 71


RESULT 14
CC238324
LOCUS       CC238324                1225 bp    DNA     linear   GSS 12-MAY-2003
DEFINITION  CH261-19I2_RM1.1 CH261 Gallus gallus genomic clone CH261-19I2,
            genomic survey sequence.
ACCESSION   CC238324
VERSION     CC238324.1  GI:30564987
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1225)
  AUTHORS   Krenitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,B. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 44
            High quality sequence stop: 100.
FEATURES             Location/Qualifiers
     source          1..1225
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-19I2"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN
1..1225

Query Match             5.8%;  Score 75.4;  DB 28;  Length 1225;
  Best Local Similarity 38.8%;  Pred. No. 0.0094;
  Matches  330;  Conservative    0;  Mismatches  512;  Indels    9;  Gaps    2;

Qy          4 AACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAACACAAAATCTTCTAACAAAAT 63
               |  |     |||  |||||| |    |||||       |   |       |||||||
Db        259 AATTTTAAAAAATAAAAATANNAANNNNNNNNNNNNNNAAAAANNNNAAATANNTNAATTATT 318
```

JOURNAL   Unpublished (2001)
COMMENT   On Feb 13, 2001 this sequence version replaced gi:12799597.
             Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0DF022BB09QP1.
FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF022YC18"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."

ORIGIN

Query Match   5.7%; Score 75; DB 9; Length 1201;
Best Local Similarity 34.5%; Pred. No. 0.011;
Matches 183; Conservative 95; Mismatches 250; Indels 3; Gaps 1;

Qy   49 ATCTTCTAACAAAATCCTAAATAAATAAGCGTTAAATTAACTAAAAAAATTAAAAAAAATG 108
          || ||||| || :| ||| || :::: |||| ::||| ||||| |:::||
Db   1184 ATATATAAYMACTAWWWTAATATWAAAATHWTHTAWAAWATWTAATAATATCAAWWWWTA 1125

Qy   109 GTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTTAGTC 168
          |:|:|| :| |||:| : |: :::| ||| :|: ||:|| | ||| :
Db   1124 TWTHTATTTTTTTTTTWWTAYWARAATAWTWWTWTTYTAAWAWAYWWTWWTTTAAWA 1065

Qy   169 ATCTTTAAGATATAAAATATCTTAAATATTCTCATGAATAAGAAAAGGAATCATCTTAAAG 228
          :|:|:||| ||||| ||| | |||| :||| ||| ::||| | :|:
Db   1064 WTATWTAWATTAAAAAAWATWTTTTATWATASWATWWAAAAWAAAWWAATWATAAW 1005

Qy   229 ACTATTAGTTGTTGAGGTACGAACATCCTTTCTTAGCATTGGGATTCTAGCTGTATGTCT 288
          : ||||| ||| ||| |: :|::|| :::| | |||| |:|::
Db   1004 WATATTTWWATTWTTTWTTTTAWATAWATWTTWTAWWWTTWATTTAYWCWTAWWTATATATTW 945

Qy   289 ATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCAGGCGAATG 348
          ||| |:|||:| | | | : ||||||| |: :|||::||::|
Db   944 TWWTWWTAWCTAATWATWWWWAAAATWWTWTAWAAAWTTATAATAAAWATAAAAWTT 885

Qy   349 GAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTACCTTCACTACAAGACTATGCC 408
          |||:| :| : |: ::::| ||| |:|||:::
Db   884 AAAAATTAAAAWAATAAATTTATWAAAAATAAWWAATATATAWWAATATWAAAWTT 825

Qy   409 AAGATTGAAGCTAGTTTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTT 468
          ||:||:| |: || : || |: || |||| |::|| | : |||
Db   824 AAAWWTTAWWATTATWATTAATMWAAWAWWAAAAWWAAAWWAATWWATWWTWWWTATAA---A 768

Qy   469 AATGCAACATTAGAACCAACTAAAAATGGCTAAAAACTAATTTAGAATCAGCCATCAACCAA 528
          ||:|| ::::||| :|||| | :::: |:
Db   767 AATTTAAAATATATAWAAAWWTATATATWAWWAAAAWAAAATWWAATTWWATTATATTWWTATAA 708

Qy   529 GCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTAGTTGAA 579
          ||||| ::: | || ||||||::||
Db   707 ATAWAWAAAAATAAAAWAWATAAWWAAAWAAATTTATATAHAHTTATATAAWWHAA 657

Search completed: May 5, 2004, 19:58:19
Job time : 2997.35 secs

RESULT 15
AL536104/c
LOCUS   AL536104   1201 bp   mRNA   linear   EST 31-MAY-2003
DEFINITION   AL536104 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
             CS0DF022YC18 5-PRIME, mRNA sequence.
ACCESSION   AL536104
VERSION   AL536104.2 GI:31260974
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 5, 2004, 09:55:26 ; Search time 11945.1 Seconds
                                   (without alignments)
                                     11571.357 Million cell updates/sec

Title: US-09-901-572A-2
Perfect score: 3189
Sequence: 1 atgaatatttctaaaaaact.........caaccgctccaaagaataa 3189

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151616995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
      10: gb_ro:*
      11: gb_sts:*
      12: gb_sy:*
      13: gb_un:*
      14: gb_vi:*
      15: em_ba:*
      16: em_fun:*
      17: em_hum:*
      18: em_in:*
      19: em_mu:*
      20: em_om:*
      21: em_or:*
      22: em_ov:*
      23: em_pat:*
      24: em_ph:*
      25: em_pl:*
      26: em_ro:*
      27: em_sts:*
      28: em_un:*
      29: em_vi:*
      30: em_htg_hum:*
      31: em_htg_inv:*
      32: em_htg_other:*
      33: em_htg_mus:*
      34: em_htg_pln:*
      35: em_htg_rod:*
      36: em_htg_mam:*
      37: em_htg_vrt:*
      38: em_sy:*
      39: em_htgo_hum:*
      40: em_htgo_mus:*
      41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3189	100.0	3189	6	AX665165	AX665165 Sequence
	2	3145.8	98.6	3189	1	AB023292	AB023292 Mycoplasm
	3	3144.2	98.6	301042	1	AB016967	AB016967 Mycoplasm
	4	3143.4	98.6	8354	1	AF214004	AF214004 Mycoplasm
	5	3143.4	98.6	10651	6	AX113685	AX113685 Sequence
	6	3057.8	95.9	3129	6	AX665242	AX665242 Sequence
	7	1052.4	33.0	7141	1	AF083976	AF083976 Mycoplasm
	8	993.2	31.1	1131	1	AB033210	AB033210 Mycoplasm
	9	924.8	29.0	1128	1	AB033211	AB033211 Mycoplasm
	10	124.6	3.9	8760	1	MYCMGP	M31431 M.genitaliu
	11	124.6	3.9	15787	1	U39698	U39698 Mycoplasma
	12	124.6	3.9	110000	6	AR300198_2	Continuation (3 of
	13	102.2	3.2	9691	1	MYCATTP	M21519 M.pneumonia
	14	102.2	3.2	16876	1	AB000002	AB000002 Mycoplasm
	15	93	2.9	1618	1	MGU34970	U34970 Mycoplasma
c	16	80.8	2.5	105682	3	AC116957_3	Continuation (4 of
c	17	80	2.5	116807	3	BX890614	BX890614 Danio rer
	18	78.4	2.5	4695	2	MGU34842	U34842 Mycoplasma
	19	76.2	2.4	146275	2	AL935272	AL935272 Danio rer
c	20	75.2	2.4	155204	2	AC007926	AC007926 Trypanoso
	21	75	2.4	213544	2	BX510939	BX510939 Danio rer
c	22	73.6	2.3	151341	5	AL929536	AL929536 Zebrafish
c	23	72.4	2.3	94534	5	AL929250	AL929250 Zebrafish
	24	71.8	2.3	64707	3	AC115607	AC115607 Dictyoste
	25	71.8	2.3	225581	3	BX537105	BX537105 Danio rer
c	26	71.4	2.3	198743	3	BX530070	BX530070 Danio rer
	27	70.8	2.2	3502	1	AY212515	AY212515 Mycoplasm
	28	70.2	2.2	183357	2	BX569779	BX569779 Danio rer
c	29	69.2	2.2	146415	5	AL929469	AL929469 Zebrafish
c	30	68.8	2.2	266544	5	AC116956	AC116956 Dictyoste
c	31	68.6	2.2	151802	3	AC114263	AC114263 Dictyoste
c	32	68.4	2.2	182676	2	BX547933	BX547933 Danio rer
	33	68	2.1	10593	3	AY232271	AY232271 Dictyoste
c	34	67.6	2.1	82139	3	AC115684	AC115684 Dictyoste
c	35	67.4	2.1	12463	3	AC115578	AC115578 Dictyoste
c	36	67.4	2.1	167913	2	BX537273	BX537273 Danio rer
c	37	66.6	2.1	111882	2	AC115612	AC115612 Dictyoste
c	38	66.4	2.1	110000	4	AC116984_3	Continuation (4 of
c	39	66.4	2.1	178664	2	BX511115	BX511115 Danio rer
	40	66	2.1	176258	2	AL929105	AL929105 Danio rer
	41	65.8	2.1	250029	2	AB014820	AB014820 Plasmodiu
c	42	65.4	2.1	169478	2	AC139280	AC139280 Homo sapi
c	43	65.4	2.1	195532	2	BX004760	BX004760 Danio rer
c	44	65.4	2.1	250195	2	AB014831	AB014831 Plasmodiu
	45	65.2	2.0	157033	2	BX323881	BX323881 Danio rer

ALIGNMENTS

RESULT 1
AX665165
LOCUS   AX665165   3189 bp   DNA   linear   PAT   26-MAR-2003
DEFINITION Sequence 2 from Patent EP1275716.
ACCESSION   AX665165
VERSION   AX665165.1   GI:29290295
KEYWORDS
SOURCE   Mycoplasma gallisepticum
  ORGANISM   Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1
  AUTHORS   Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki,Y.
  TITLE   Modified dna molecule, recombinant containing the same thing, and
            uses thereof
  JOURNAL   Patent: EP 1275716-A 2 15-JAN-2003;

Pred. No. is the number of results predicted by chance to have a

FEATURES   Zeon Corporation (JP)
source   Location/Qualifiers
             1..3189
             /organism="Mycoplasma gallisepticum"
             /mol_type="unassigned DNA"
             /db_xref="taxon:2096"
             /note="mgc3 gene"
ORIGIN

Query Match   100.0%;   Score 3189;   DB 6;   Length 3189;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 3189;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA   60

Qy    61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120

Qy   121 TTAGTTAATCAAGCAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGTCTTGGACAA   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTAGTTAATCAAGCAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGTCTTGGACAA   180

Qy   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240

Qy   241 AATGGAACAATTATCAAATTAGATAGTTTACTAAAACCATTATATGGTTTAGATCTAAGT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AATGGAACAATTATCAAATTAGATAGTTTACTAAAACCATTATATGGTTTAGATCTAAGT  300

Qy   301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACAACACTAGCAGAAAT  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACAACACTAGCAGAAAT  360

Qy   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCTGTTGGTTAATGATGAAGCTAAC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCTGTTGGTTAATGATGAAGCTAAC  420

Qy   421 GTTCATTTAAAAGAATTAATACTAACTAACTAACTAATGGTAATGAATTAGAACAACAATTCT  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTTCATTTAAAAGAATTAATACTAACTAACTAACTAATGGTAATGAATTAGAACAACAATTCT  480

Qy   481 AAGTTGTAATTGGTGGTGTTGTTCAAACCAAACTCAGGTGAAATTGTTAATGACTTCATTTGAGAT  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGTTGTAATTGGTGGTGTTGTTCAAACCAAACTCAGGTGAAATTGTTAATGACTTCATTTGAGAT  540

Qy   541 ACTAAATTTAATTTACAAAACCACTACTACTAATCCTTCAAATTCCTTCAAACTTTATACATTCAACTTTATACATT  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ACTAAATTTAATTTACAAAACCACTACTACTAATCCTTCAAATTCCTTCAAACTTTATACATTCAACTTTATACATT  600

Qy   601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTGAGTTAGTGGA  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTGAGTTAGTGGA  720

Qy   721 ACAAATGCTGATGATGGGATGTTTGATTGTGTGGAATGGTCAAATAACTAATAACAGATCCT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAAATGCTGATGATGGGATGTTTGATTGTGTGGAATGGTCAAATAACTAATAACAGATCCT  780

Qy   781 ATTGCTCAAACTAAAAACCACTACTACTAGATGACGATTCTCAACTTTAATTCAGGAGCA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATTGCTCAAACTAAAAACCACTACTACTAGATGACGATTCTCAACTTTAATTCAGGAGCA  840

Qy   841 ATGCCTGGTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGTCAGCATAGAATTAAAACA  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATGCCTGGTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGTCAGCATAGAATTAAAACA  900
```

```
Qy   901 TCTTTCCAATTAGATGATGAAAAATTGTTTATCCAGAATGGACTGGTTCTTGAACGAATAAA   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTCCAATTAGATGATGAAAAATTGTTTATCCAGAATGGACTGGTTCTTGAACGAATAAA   960

Qy   961 AATATTACAAGATTAGCTACTGGAAGTTGCAAGCAACGAACAAGATATTGGATTCTTGAC  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AATATTACAAGATTAGCTACTGGAAGTTGCAAGCAACGAACAAGATATTGGATTCTTGAC  1020

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTGACGTATTTCAAGACTA  1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTGACGTATTTCAAGACTA  1080

Qy  1081 TACTTAAACTCAGTAATTCTTTATCATTCATTGGTGATAGTATTTATATTTTTGGTACC  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TACTTAAACTCAGTAATTCTTTATCATTCATTGGTGATAGTATTTATATTTTTGGTACC  1140

Qy  1141 TCTGAATTACCATCATTATGGTACTATCATTCCCAACTAGATATCGATCTAACCGCT  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TCTGAATTACCATCATTATGGTACTATCATTCCCAACTAGATATCGATCTAACCGCT  1200

Qy  1201 TTGAATCAAGTTAAAAACAGATGATATATGAAGCTTCAAGCACTGATAACGGTACAACACA  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TTGAATCAAGTTAAAAACAGATGATATATGAAGCTTCAAGCACTGATAACGGTACAACACA  1260

Qy  1261 AACGGAACAACGACAACAGCTGATACATCTCAGTGGTTCAACAGGTGCGTGGGAACAGGAAAT  1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AACGGAACAACGACAACAGCTGATACATCTCAGTGGTTCAACAGGTGCGTGGGAACAGGAAAT  1320

Qy  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTACTTAAATACTTAAATACTTCGTAGTTTTGGA  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTACTTAAATACTTAAATACTTCGTAGTTTTGGA  1380

Qy  1381 ATTGATAGTAACAACCAACTTCTGCAAACAAAATAGTGAAACTAATTGGGCAGATCCTAAC  1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 ATTGATAGTAACAACCAACTTCTGCAAACAAAATAGTGAAACTAATTGGGCAGATCCTAAC  1440

Qy  1441 GTTATTGAAGCAAGAATATATGCTGAATACAGATTACAGATTAGCTATTCAAAATGAAATTCCAATA  1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTTATTGAAGCAAGAATATATGCTGAATACAGATTACAGATTAGCTATTCAAAATGAAATTCCAATA  1500

Qy  1501 ACTAATGCAGGAAACTTTATCCGAAACAACAATTGGTGTGTTGGTTTTACTTCGGAAACTTCAA  1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ACTAATGCAGGAAACTTTATCCGAAACAACAATTGGTGTGTTGGTTTTACTTCGGAAACTTCAA  1560

Qy  1561 TCAAGAGTAGTTTTTAAGAGCTTCATTGGTTATTTAGGATACCAACAAACTAGAACACAGGAACTTCTGG  1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TCAAGAGTAGTTTTTAAGAGCTTCATTGGTTATTTAGGATACCAACAAACTAGAACACAGGAACTTCTGG  1620

Qy  1621 CCTTTCTTATACGTATTTGGTTATTGGTTTTAAACAACAACAGCCCCTTACGACGTATTAGATTCTCCAAGAGTA  1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CCTTTCTTATACGTATTTGGTTATTGGTTTTAAACAACAACAGCCCCTTACGACGTATTAGATTCTCCAAGAGTA  1680

Qy  1681 TACGGAACATATAAGCTTTAAACAACAGCCCCTTACGACGTATTAGATTCTCCAAGAAAG  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACGGAACATATAAGCTTTAAACAACAGCCCCTTACGACGTATTAGATTCTCCAAGAAAG  1740

Qy  1741 GGTACTGAAAACCAATCAATTAGAGAACTTCATTACATCCCTGTATGGGTGGATAT  1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 GGTACTGAAAACCAATCAATTAGAGAACTTCATTACATCCCTGTATGGGTGGATAT  1800

Qy  1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATAGAGCACAAGGTGAC  1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATAGAGCACAAGGTGAC  1860

Qy  1861 ACACCAGAAAGCCGAAGCATCTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT  1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACACCAGAAAGCCGAAGCATCTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT  1920

Qy  1921 CAATCAGTTTAGGATTTGATGAATTAGAAAATAACTTAAATGTTGGGGTTAAAGCATCA  1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CAATCAGTTTAGGATTTGATGAATTAGAAAATAACTTAAATGTTGGGGTTAAAGCATCA  1980

Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACACA  2040
```

```
Db    3061  GCTCCAGCTAAACCATCTGACCAACGAACGTAGCTCACCAGCTAAACCAACCAACGGCTAAA 3120
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    3121  TCTGGTGCGGCTACAAAAACCAACTGCTCCTAAGCCAGTCGCTGCTCCAAAAACCAACCGCTCC 3180
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121  TCTGGTGCGGCTACAAAAACCAACTGCTCCTAAGCCAGTCGCTGCTCCAAAAACCAACCGCTCC 3180

Qy    3181  AAAGAATAA 3189
                   |||||||||
Db    3181  AAAGAATAA 3189


RESULT 2
AB023292
LOCUS       AB023292                3189 bp    DNA    linear   BCT 27-MAY-2000
DEFINITION  Mycoplasma gallisepticum gene for 120kDa membrane protein, complete
            cds.
ACCESSION   AB023292
VERSION     AB023292.1  GI:4240363
KEYWORDS    120kDa membrane protein.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (sites)
  AUTHORS   Yoshida,S., Fujisawa,A., Tsuzaki,Y. and Saitoh,S.
  TITLE     Identification and expression of a Mycoplasma gallisepticum surface
            antigen recognized by a monoclonal antibody capable of inhibiting
            both growth and metabolism
  JOURNAL   Infect. Immun. 68 (6), 3186-3192 (2000)
   MEDLINE  20278096
   PUBMED   10816462
REFERENCE   2  (bases 1 to 3189)
  AUTHORS   Yoshida,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1999) Shigeto Yoshida, Jichi Medical School,
            Department of Medical Zoology; Yakushiji 3311-1,
            Minamikawachimachi, Tochigi 329-0498, Japan
            (E-mail:shigeto@jichi.ac.jp, Tel:81-285-58-7339,
            Fax:81-285-44-6489)
FEATURES             Location/Qualifiers
     source          1..3189
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:2096"
     gene            1..3189
                     /gene="M11"
     CDS             1..3189
                     /gene="M11"
                     /codon_start=1
                     /transl_table=4
                     /product="120kDa membrane protein"
                     /protein_id="BAA74942.1"
                     /db_xref="GI:4240364"
                     /translation="MNISKKLKSYTLIGGLAVFGALGSASFGFKQSDKSNDNTQLVNQ
                     ARTLDANSVRLAGLGQNGSLFNVLRDVDDNFITAANGTIIKLDSFTKPLYGLDLSDD
                     FAGYKVKQIVSDYTTSRNFDDRQTRAYYALLVNDEANVHLKRINTNSNRIGNRNNNS
                     KFVIGGVDNPAHVIRFTDDGTKFNFTKQTGEIVNDFILDAPILPKDLHPDWYNLYIQ
                     RKLPDVNTAVVPWPVGRVSGTNADDGMPDFGNGQITNTDPIAQTKTTDNQNPSTF
                     NSGAMPGANNRYDSQLNVKGHRIKTSFQLDEKFVYPEWTGSSENKNITRLATGSLPSNE
                     RYWILDIPGTQVTLKEDSVNVFSRLYLNSVNSLSFIGDSIYLFGTSELPSLWYYSFP
                     TRLSDLTALNQVKTDDIEASSTDNGTTVTNGTTTTADTSSGSTGAGTGNTTNTSQTVSN
                     PTLNTYRSFGIDSKPTSANKIDETNWADPNVIEARIYABYRLGIQNEIPITNAGNFIR
                     NTIGGVPFTSTGSRVVLRASYNGDQREPTNGFPLYVPGYLVGYQQTRTGTFWYGTYKL
                     LNNSPYDVLDAARVGETENQFRRTSLTYPVMGGYLTEBGARSPSNTPYIRAQGDTPES
                     RSIFQSGYSDNTYBYIQSVLGFDGIRNNLNVGVKASSFLNSNRPNPNGLEMIAATTYL
                     RSCIGLARTSGLPENQQPFGTTHQVISVPGDQFSSIKNIRTIFPGNQLMYFLFTNENN
                     KSSVYTLRIADSSNPDASSFSPTSLIIDVNEIGVILPLLDNSFYTVNAAGNVALFSSN
                     PGSPGSYTAVNTFNQNLSDIAFBGSGAKYTSDFWGTIQFKPDEYLIQNGFTSQVARNF
                     VTNQSPLNSLVDFTPANAGTNRYRVVDPDGNLTNQNLPLKVQIQYLDGKYVDAKLKNN
                     NLVTFSYNNFAALPSWVVPTAIGSTLGILAIMILGLAIGIFLRAQRKLQDKGFKTTF
                     KVDTLTAAVGSVYKKIITQTANVKKKPAALGAGKSGDKKPAAAAKPAAPAKPSAPKA
                     SSPAKPTAPKSGAPTKPTAPKPAAPKPTAPKE"
ORIGIN
```

```
Query Match         98.6%;  Score 3145.8;  DB 1;  Length 3189;
Best Local Similarity  99.2%;  Pred. No. 0;
Matches 3162;  Conservative 0;  Mismatches 27;  Indels 0;  Gaps 0;

Qy     1 ATGAATATATTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAATATATTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGA   60

Qy    61 GCTCTTGGTTCTGCAAGCTTGGCTTTGGCTTTAAGCAATCAGATAGAGTAACGATAACACCGAA  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCTCTTGGTTCTGCAAGCTTGGCTTTGGCTTTAAGCAATCAGATAGAGTAACGATAACACGCAA  120

Qy   121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAAATTCTGTTGAGCTTCAGGTCTTGGACAA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAAATTCTGTTGAGCTTCAGGTCTTGGACAA  180

Qy   181 AATGGTTCGTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATGGTTCGTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCT  240

Qy   241 AATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATGGTTTAGATCTAAGT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATGGTTTAGATCTAAGT  300

Qy   301 GATGATTGGTGATCAAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GATGATTTGCTGGATACAAAGTAAAACAAACAAGAGTTTCAGATTACACAACTAGCAGAAAT  360

Qy   361 AGATTTGATCAAAGACAAAGAGACAATATTATGCTCGTTGGTTAATGATGAAGCTAAC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGATTTGATCAAAGACAAAGAGACAATATTATGCTCGTTGGTTAATGATGAAGCTAAC  420

Qy   421 GTTCATTTAAAAAAGAATTAATACTAACTCAAATAGAATTGTAATAGAACAACAATTCT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTTCATTTAAAAAAGAATTAATACTAACTCAAATAGAATTGTAATAGAACAACAATTCT  480

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTTACTACTGATGATGGG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTTACTACTGATGATGGG  540

Qy   541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600

Qy   601 GCGCCAATCTTACCTAAAGATTACACCGCAGTTGGTATAACTATACATTTCAAAGAAAG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCGCCAATCTTACCTAAAGATTACACCGCAGTTGGTATAACTATACATTTCAAAGAAAG  660

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGTTAGTTGGA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCTTACCAAATGACGTAAACAACTGCAGTTGTTCCTTGGCCAGTAGTTAGTTGGA  720

Qy   721 ACAAATGCTGATGATGGGATGTTTGATTGTGGGAAGATGTCAAATAACTAATACAGATCCT  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAAATGCTGATGATGGGATGTTTGATTGTGGGAAGATGTCAAATAACTAATACAGATCCT  780

Qy   781 ATTGCTCAAACTACTAACGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATTGCTCAAACTACTAACGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA  840

Qy   841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGGTGGTTAAATGTGGGTTAAAGCATCA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGGTGGTTAAATGTGGGTTAAAGCATCA  900

Qy   901 TCTTTCCAATTGAGTGATGAAAATTGTTTTATCCAGAAATGGACTGGTTCTGAAGAGAATAAA  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTCCAATTGAGTGATGAAAATTGTTTTATCCAGAAATGAACTGGTTCTGAAGAGAATAAA  960

Qy   961 AATATTACAAGATCACAAGTTGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AATATTACAAGATCACAAGTTGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020


Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTGTGTATTTCAAGACTA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ATACCAGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTGTGTATTTCAAGACTA 1080

Qy  1081 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTAGTATTTATATTTTGGTACC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTAGTATTTATATTTTGGTACT 1140

Qy  1141 TCTGAATTACCATCATTATGTGTACTATTCATTCCAACTAGATTATCTGATCTAACCGCT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TCTGAATTACCATCATTATGTGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT 1200

Qy  1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260

Qy  1261 AACGGAACAACAACACAGCTGATACATCTGATACAGTTCAACAGGTGCTGGAACAGAGAAAT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AACGGAACAACGACACAGCTGATACATCTGATACAGTTCAACAGGTGCTGGAACACAGGAAAT 1320

Qy  1321 ACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTAAATACTTATCGTAGTTTTGGA 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTAAATACTTATCGTAGTTTTGGA 1380

Qy  1381 ATTGATGATAAAACCACTTCTGCAACAAAATAGTTTCAGATTACACAACTAGCAGATCCTAAC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 ATTGATAGTAGTAAACCACTTCTGCAACAAAATAGTTTCAGATTACACAACTAGCAGATCCTAAT 1440

Qy  1441 GTTATTGAAGCAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAAATTCCAATA 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTTATTGAAGCAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAAATTCCAATA 1500

Qy  1501 ACTAATGCAGGAAAACTTATCTCCGAAACACAATCAATTGGTGGTTGGTTTACTTCAACAGGT 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ACTAATGCAGGAAAACTTATCTCCGAAACACAATCAATTGGTGGTTGGTTTACTTCAACAGGT 1560

Qy  1561 TCAAGAGTAGTTTAAGAGCTTCTTATATAGAGAGTCGATCAACGTGATCAACTGGAAACTTCCAA 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TCAAGAGTAGTTTAAGAGCTTCTTATATAGAGAGTCGATCAACGTGATCAACTGGAAACTTCCAA 1620

Qy  1621 CCTTTCTTATACGTATTTGGTTATTGGTTATTGGTATTTAAACACAGCCCTTACGACGTATTAGAT 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CCTTTCTTATACGTATTTGGTTATTTGGTTATTGGTATTTAAACACAGCCCTTACGACGTATTAGAT 1680

Qy  1681 TACGGAACATATAAGCTTTAAACACTATTAAACAACAGGCTATTCCCAAGAGTA 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACGGAACACATATAAACTATTAAACAACAGGCTATTCCCAAGAGTA 1740

Qy  1741 GGTACTGAAACCAATCAATTTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATAT 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 GGTACTGAAACCAATCAATTTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATAT 1800

Qy  1801 CTAACTGAAGAGGTGCTAGAAGTTCTCTTAATACTAACTAATACAGATCCT 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CTAACTGAAGAGGTGCTAGAAGTTCTCTTAATACTAACTAATACAGATCCT 1860

Qy  1861 ACACCAGAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACACCAGAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT 1920

Qy  1921 CAATCAGTTTAGGATTTGATGGAATTAGAATAAACTAAATGTTGGGTTAAAGCATCA 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CAATCAGTTTAGGATTTGATGGAATTAGAATAAACTAAATGTTGGGTTAAAGCATCA 1980

Qy  1981 AGCTTCTTAAACTCAAATGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 AGCTTCTTAAACTCAAATGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA 2040

Qy  2041 TACTTAAGATCACCAAGTTGATTAGCTAGACATCTGGATTACCAACCAACATTC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TACTTAAGATCACAAAGTTGATTAGCTAGACATCTGGATTACCAACCAACATTC 2100

Qy  2101 GGAACAACTCACCAAGTTATTTCAGTATCACCTGGTGATCATCAATTAAGAAT 2160
```

Db   3181 AAAGAATAA 3189

```
RESULT 3
Db       AB016967
LOCUS       AB016967                301042 bp    DNA     linear   BCT 03-SEP-2003
DEFINITION  Mycoplasma gallisepticum strain R section 1 of 4 of the complete
            genome.
ACCESSION   AB016967 AB015450
VERSION     AB016967.1  GI:31541048
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum R
  ORGANISM  Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 301042)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE     The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
  PUBMED    12949158
REFERENCE   2  (bases 1 to 301042)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzel,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathobiology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eaglevile Road U-89, Storrs, CT
            06269-3089, USA
FEATURES             Location/Qualifiers
     source          1..301042
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:233150"
     rep_origin      1..1889
                     /note="putative oriC"
     protein_bind    complement(24..32)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence ttatcMaMa"
                     pulmonis consensus ttatcMaMa"
                     /bound_moiety="DnaA"
     protein_bind    complement(36..44)
                     /function="initiation of the chromosome replication"
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence ttatcMaMa"
                     pulmonis consensus ttatcMaMa"
                     /bound_moiety="DnaA"
     protein_bind    114..122
                     /function="initiation of the chromosome replication"
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence ttatcMaMa"
                     /bound_moiety="DnaA"
     repeat_region   209..217
                     /note="oriC imperfect repeat"
     protein_bind    complement(254..262)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence YtWtMHaMR"
                     mycoides consensus YtWtMHaMR"
                     /bound_moiety="DnaA"
     protein_bind    298..306
                     /function="initiation of the chromosome replication"
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence YtWtMHaMR"
                     mycoides consensus YtWtMHaMR"
                     /bound_moiety="DnaA"
     repeat_region   complement(401..450)
                     /function="initiation of the chromosome replication"
                     /note="oriC imperfect repeat"
     repeat_region   complement(462..510)
                     /note="oriC imperfect repeat"
```

```
Db   2101 GGAACAACTCACCAAGTTATTTCAGTATCAGTATCACCTGGTGTATCATCCTGGTGATCAGTTCTCATCATTCAATTAAGAAT 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2161 ATTAGAAACAATCTTCCCTGGTAACCAGTTATGTACTTCTTATTCACAAATGAAAATAAT 2220

Db   2161 ATTAGAAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCACAAATGAAAATAAT 2220
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Qy   2221 AAATCTAGTGTTTATACATTAAGATAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280

Db   2221 AAATCTAGTGTTTATACATTAAGATAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGAC 2340

Db   2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGAC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2341 AATTCATTCTATACAGTAAATGCTGCTGATGTTGCATTGTTCTCATCAAACCCTGGT 2400

Db   2341 AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGGT 2400
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy   2401 TCTCCTGGATCATATACGTCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460

Db   2401 TCTCCTGGATCATATACGTCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTCTGGGGAACAATCCAATTCAAACCCGAT 2520

Db   2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTCTGGGGAACAATCCAATTCAAACCCGAT 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTACTGGCTAGAAACTTCGTTACAAACCAA 2580

Db   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTACTGGCTAGAAACTTCGTTACAAACCAA 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2581 AGCTTCTTAAAACAGTTAGTTGACTTCACTCCGTACTAATGCTGGTACTAACTACCGTGTA 2640

Db   2581 AGCTTCTTAAAACAGTTAGTTGACTTCACTCCGTACTAATGCTGGTACTAACTACCGTGTA 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2641 GTGGTTGATCCTGATGGGTAATTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700

Db   2641 GTGGTTGATCCTGATGGGTAATTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2701 TACTTAGATGGTAAGTAAGTATATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 2760

Db   2701 TACTTAGATGGTAAGTAAGTATATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2761 TATAACAACTTTGGCGCCTTACCTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820

Db   2761 TATAACAACTTTGGCGCCTTACCTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA 2880

Db   2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA 2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940

Db   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2941 GCTGCTTTAGGTGCTGGTAAATCTGGTAATAAGAAACCTGCTAACCTGCTAAACCTGCT 3000

Db   2941 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTAACCTGCTAAACCTGCT 3000
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy   3001 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTAACCTGCTAAACCTGCT 3060

Db   3001 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTAACCTGCTAAACCTGCT 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA 3120

Db   3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3121 TCTGGTGCGCGCCTACAAAACCAACTGCTCCTCCAAAACCAACCAACCGGCTCCC 3180

Db   3121 TCTGGTGCGCGCCTACAAAACCAACTGCTCCTCCAAAACCAACCAACCGGCTCCC 3180

Qy   3181 AAAGAATAA 3189
          |||||||||
```

```
repeat_region   complement(575..623)
                /note="oriC imperfect repeat"
repeat_region   complement(713..761)
                /note="oriC imperfect repeat"
repeat_region   complement(763..811)
                /note="oriC imperfect repeat"
repeat_region   complement(821..869)
                /note="oriC imperfect repeat"
repeat_region   872..920
                /note="oriC imperfect repeat"
repeat_region   991..1004
                /note="oriC imperfect repeat"
repeat_region   1061..1109
                /note="oriC imperfect repeat"
repeat_region   1147..1195
                /note="oriC imperfect repeat"
protein_bind    complement(1185..1192)
                /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                binding site; consensus sequence approach; Mycoplasma
                capricolum consensus ttwtMtaa"
                /bound_moiety="DnaA"
                /function="initiation of the chromosome replication"
repeat_region   complement(1230..1278)
                /note="oriC imperfect repeat"
repeat_region   1326..1374
                /note="oriC imperfect repeat"
repeat_region   complement(1551..1564)
                /note="oriC imperfect repeat"
repeat_region   complement(1609..1657)
                /note="oriC imperfect repeat"
repeat_region   complement(1700..1748)
                /note="oriC imperfect repeat"
repeat_region   complement(1831..1879)
                /note="oriC imperfect repeat"
gene            1890..2681
                /gene="parA"
                /locus_tag="MGA_0619"
                /synonym: Soj"
CDS             1890..2681
                /gene="parA"
                /locus_tag="MGA_0619"
                /note="MGA_0619 ATPase involved in chromosome
                partitioning; pfam00991 ParA family ATPase; COG1192 Soj
                ATPases involved in chromosome partitioning, MGR_001"
                /codon_start=1
                /transl_table=4
                /product="ParA/Soj"
                /protein_id="AAP56351.1"
                /db_xref="GI:31541049"
                /translation="MMIISFINNKGGVLKTTLATNICGVFSKFFPKSRSVIVDLDGQG
                NVSATFGQHPERLKNTLLDIFRGEKDIDDCVLNVFPSIDILPSNHELSFVDMDVARKE
                YKLSVIKNLIEKLEEMYDPVFLDIPPAMSTIVSVSMHISDMIVIPFEPDQYSILGLLR
                VIDTIETFPERNEDLKVLVVPTKVNARTRLHNDVLNIVKTKLSKKNIPLSKNIVSFTT
                KSSASVGYEKLPIVLINQRSKYQEEYIEITKEIINLLKDNKNNEK"
gene            2659..3147
                /locus_tag="MGA_0621"
CDS             2659..3147
                /locus_tag="MGA_0621 PS503318; MGR_002"
                /codon_start=1
                /transl_table=4
                /product="unique hypothetical"
                /protein_id="AAP56352.1"
                /db_xref="GI:31541050"
                /translation="MTRTMKNKAKKKERRFTDLSADLDEEVEKIDPEYEDRKEIKIE
                KNKDNQVIDKNDPFYSESFEEARIQLIKDKKVEVKKEEKVQETTVKNKISEAKKEE
                AKDVYIDSSLEIASQEPLTKGMHFYTNSRIIRKVRECAKNKGLSISRLITMILDKSIK
                EE"
gene            3163..4548
                /gene="dnaA"
                /locus_tag="MGA_0622"
CDS             3163..4548
                /gene="dnaA"
```

```
                /locus_tag="MGA_0622"
                /note="MGA_0622 ATPase involved in DNA replication
                initiation COG0593"
                /codon_start=1
                /transl_table=4
                /product="DnaA"
                /protein_id="AAP56353.1"
                /db_xref="GI:31541051"
                /translation="MKTKLKRFLEISVHFNEANSELLDAFVHSIDFVFEENDNIYIY
                PESPYFFNEFKNKLNHLINVENAVVFNDYLSLEWKCIIKENKRVNLLNKKEADTLKEK
                LATLKKQEKYKINPLSKGIKEKYNFGNYILVFEFNKEAVYLAKQIANKTTHSNWNPIII
                EGKPGVGKSHLLQAIANERQKLPPEKICVLSSDDFGSBFLKSVIAPDPTHIESFKSK
                YKDYDLLMIDDVQIISNRPKTNETPFTIFNSLVDQKKTIVITLDCKIBEIQDKLTARM
                ISRFQKGINVRINQPNKNEIIQIFKQKFKENNLEKYMDDHVIEEISDFDEGDIRKIEG
                SVSTLVFMNQMYGSTKTKDQLLKSFIEKVTNRKNJILSKDPKYVFDKIKYHFNVSEDV
                LKSSKRKKEIVQARHICMVYLKNVYNKNLSQIGKLLRKDHTTVRHGIDKVEBELENDP
                NLKSFLDLFKN"
gene            4744..5133
                /locus_tag="MGA_1322d"
                /note="MGA_1322d"
CDS             4744..5133
                /locus_tag="MGA_1322d"
                /note="Q COG1132 ABC-type multidrug/protein/lipid
                transport system, ATPase component; similar to MGA_1284;
                MGR_004"
                /codon_start=1
                /transl_table=4
                /product="ABC transporter component"
                /protein_id="AAP56354.1"
                /db_xref="GI:31541052"
                /translation="MNRKLDHCISALLYFGLINFGIIGRDKNALTQYSLILLGVAAVG
                PTFGYIGGRSIIASVBFAKQLRVNIFERYQSFSVKNTDKFEKASVLTRMTTDINFIH
                QSIQSGRTAIRGMSVPLFSLVLMFVTS"
gene            5130..6077
                /locus_tag="MGA_0625"
                /note="synonym: mldB"
CDS             5130..6077
                /locus_tag="MGA_0625"
                /note="MGA_0625, MdlB-like; ABC-type
                multidrug/protein/lipid transport system MdlB [Q] COG1132
                PS00890; similar to MGA_1285, MGA_0626 and MGA_1287;
                MGR_005"
                /codon_start=1
                /transl_table=4
                /product="MdlB-like"
                /protein_id="AAP56355.1"
                /db_xref="GI:31541053"
                /translation="MKLGIASVAIMPVIIGGILLVYRFVIGNYKKLPKQYDQLNNLAK
                ESIAGARVVKSTHQQDNEIQKFNRVAGFIYKNFTKIRITALISPIVLFCIYALAIAI
                AWIGTNNIVDGKLDIGSLASVPAYAFQMLINLLLSVVYVTIITAKPSKDRIIEVLTE
                KIDIKDKGVIDDVSDYEDVSFKVDNPHHNLEKINIKIKKGQTIGIGSTGS
                GKTSIVNLLTRLYECNEBGQVLLNNIQLNNYSIKALRDAIAIVPQKSILYSGTIKDNIL
                MGGNYSDEEVEKAITQAQAEFINKLPNKLDSIVEQKWN"
gene            6369..8294
                /locus_tag="MGA_0626"
CDS             6369..8294
                /locus_tag="MGA_0626"
                /note="MGA_0626, ABC-type multidrug/protein/lipid
                transport system; similar to MdlB [Q] COG1132 PS00890;"
```

```
Query Match             98.6%;  Score 3144.2;  DB 1;  Length 301042;
Best Local Similarity   99.1%;  Pred. No. 0;
Matches 3161;  Conservative 0;  Mismatches 28;  Indels 0;  Gaps 0;

Qy        1 ATGAATATTTCTAAAAAACTAAAAGTTATACATTGATAGGTCGATGGATTAGCTGTATTTCGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226607 ATGAATATTTCTAAAAAACTAAAAGTTATACATTGATAGGTCGATGGATTAGCTGTATTTCGA 226666

Qy       61 GCTCTTGGTTCTGCAAGCTTGGCTTTAAGCAATCAGATCAGATAAGAGTAACCATAACACGCAA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226667 GCTCTTGGTTCTGCAAGCTTGGCTTTAAGCAATCAGATCAGATAAGAGTAACGATAACACGCCA 226726

Qy      121 TTAGTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTCGTTAGACTCTTGGACAA 180
```

REFERENCE 1 (bases 1 to 8354)
AUTHORS   Goh,M.S., Gorton,T.S., Forsyth,M.H., Troy,K.E. and Geary,S.J.
TITLE   Molecular and biochemical analysis of a 105 kDa Mycoplasma
            gallisepticum cytadhesin (GapA)
JOURNAL   Microbiology 144 (Pt 11), 2971-2978 (1998)
MEDLINE   99061189
PUBMED   9846732
REFERENCE 2 (bases 1 to 8354)
AUTHORS   Papazisi,L., Troy,K.E., Gorton,T.S., Liao,X. and Geary,S.J.
TITLE   Analysis of cytadherence-deficient, GapA-negative Mycoplasma
            gallisepticum strain R
JOURNAL   Infect. Immun. 68 (12), 6643-6649 (2000)
MEDLINE   20536401
PUBMED   11083776
REFERENCE 3 (bases 1 to 8354)
AUTHORS   Papazisi,L., Troy,K.E. and Geary,S.J.
TITLE   Direct Submission
JOURNAL   Submitted (09-DEC-1999) Department of Pathobiology, University of
            Connecticut, 61 North Eagleville Rd. U-89, Storrs, CT 06268-3089,
            USA
FEATURES   Location/Qualifiers
    source   1..8354
                /organism="Mycoplasma gallisepticum"
                /mol_type="genomic DNA"
                /strain="R"
                /db_xref="taxon:2096"
    gene   667..3756
                /gene="gapA"
    CDS   667..3756
                /gene="gapA"
                /note="GapA; MGP1; cytadhesin; similar to Mycoplasma
                pneumoniae P1"
                /codon_start=1
                /transl_table=4
                /product="adherence protein A"
                /protein_id="AAF25381.1"
                /db_xref="GI:6694735"
                /translation="MIWTFDPGNTNDLTGKVGFYDANNRLTAPSGDVPFNVSDLSSKT
                VVEATDQEDPNVFYLLLIPDAAVQQEQKTKDQVPENYFMSDAPATGDTSAEGSATPA
                GGGSSSAACGGAVAPAAASSTARLVBEGNSAGMGTMTPTASTSETVIDYNSDQNKIP
                KPKTLLDDSSESSESINGGRTYANINTQNNLQGVIVKVNENLFNSENPFAVENMARIKP
                KDMVDNYPSTWTQGSANGKMTNVLQCYKHDNPAVNNRPYRAKYYPKRLETQTTTPLI
                DSSFSPYEHPEWYEDNQFVMPWMQYITNLGGLYAKDGMVYLFGGNGTWVNNESALSIG
                VFRTKFENRIAEAPGNTKTVGYPYGILLSAISFDATRNGLALAPASLGQDVNYTVPR
                LAVGGVSSPRGANGNIFLGSAITWGTNGGNFLDTKWHSPAVIEDAPTTFVTVNSSGAL
                QNSGNPQPTSTPMPNSNGNESIPYRWTNSYDYNSVRFPAALISKPAGGNTKQVESLFTT
                ALKLDTLNSLPNKFTQENNIFFSYAMLDGRQWSLGTRKOSAWLATNTINNPTVNTQQQ
                LASTVAGENANPRNIINALTTAKGFDRRDIGNVVTYSNNTNKFTYYQVGGAITTWP
                EVQVNYKTSANITYNLTRTDFGSTTPATQDANTVSSKLNGAYLSSTGDQQGWYNGSI
                YVKKASFTPSSQGYTWQDFKGLTTTASNAVISNWTKAGYSIRPDDDTVFNASKIPPEK
                EITAAVNVRSLDSYYVQLNGETSVNTVARVSPDSSALALNPNRITNPLMNRDNVIGQG
                AFISRNDIPSSFPFENKINDIVTTEADGKEVLDSKYINSIYRYTPQNNPDIRLRLLVI
                DRSRATNDPIKLLPQVLVDGEYVAVPQANSVFVSDQBFTGFDALPGYVLPVAISIPII
                IIALALALGLGIGIPMSQNRKMLKQGFAISNKKVDILTTAVGSVFKQIINRTSVTNIK
                KTPQMLQANKDGASSPSKPSAPAKKPAGPTKSAPGAKPTAPAKPKPAPTKKIE"
    gene   3778..6966
                /gene="crmA"
    CDS   3778..6966
                /gene="crmA"
                /note="CrmA; similar to Mycoplasma pneumoniae MGPC,
                Mycoplasma genitalium MGPC, Mycoplasma galliseptium
                strain S6 GapA, Mycoplasma pirum P1-like adhesin,
                Mycoplasma pneumoniae ADP1, and Mycoplasma genitalium
                ADP1, encoded by GenBank Accession numbers Q50341, P22747,
                AF083376, Q50371, AAB95661 and P20796, respectively"
                /codon_start=1
                /transl_table=4
                /product="putative cytadherence related molecule A"
                /protein_id="AAF25382.1"
                /db_xref="GI:6694736"
                /translation="MNISKKLKSYTHLIGGLAVFGALGSASPGFKQSDKSNDNTQLVNQ
                ARTLDANSVRLAGLGQNGSLPNTVLRDVDVDDNFITRANGTIIKLDSFTKPLYGLDLSDD

Qy   2341 AATTCATTCTATACAGTAAATGCTGCTGGTAAATGTTGCATTGTTCTCATCATCAAACCCTGGT 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 228947 AATTCATTCTATACAGTAAATGCTGCTGGTAAATGTTGCATTGTTCTCATCATCAAACCCTGGT 229006

Qy   2401 TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229007 TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 229066

Qy   2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGAACAATCCAATTCAAACCCGAT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229067 GAAGGTTCTGGTGCTAAATATACATCTGATTTCTGAGGAACAATCCAATTCAAACCCGAT 229126

Qy   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229127 GAGTACTTAATTCAAAATGGGTTCACTAGTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 229186

Qy   2581 AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229187 AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 229246

Qy   2641 GTGGTTGATCCTGATGGTAATTAACAAACCAAAACTACCTCTAAAGTTCAGATCCAA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229247 GTGGTTGATCCTGATGGTAATTAACAAACCAAAACTACCTCTAAAGTTCAGATCCAA 229306

Qy   2701 TACTTAGATGTGTAAGTATATTATGATGCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229307 TACTTAGATGTGTAAGTATATTATGATGCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 229366

Qy   2761 TATAACAACTTTGGCGCCTTACCTTCATGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229367 TATAACAACTTTGGCGCCTTACCTTCATGGTAGTGCCTACAGCAATTGGTAGTACATTA 229426

Qy   2821 GGTATTCTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229427 GGTATTCTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA 229486

Qy   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229487 AGAAAATTACAAGACAAAGGGTTCAAAACAACACATCAAAAAAGTTGATACCTTGACTGCT 229546

Qy   2941 GCTGTTGGTTCAGTTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229547 GCTGTTGGTTCAGTTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 229606

Qy   3001 GCTGCTTTAGTGCTGGTAAATCTGGTAGATAAGAAACCTGCTGCTGTGCTAAACCTGCT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229607 GCTGCTTTAGTGCTGGTAAATCTGGTAGATAAGAAACCTGCTGCTGTGCTAAACCTGCT 229666

Qy   3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229667 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA 229726

Qy   3121 TCTGGTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCAGCTGCTCCAAAACCAACCGCTCCC 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229727 TCTGGTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCAGCTGCTCCAAAACCAACCGCTCCC 229786

Qy   3181 AAAGAATAA 3189
            |||||||||
Db 229787 AAAGAATAA 229795

PAGYKVKQIVSDYTTSRNRFDQRQTRAYYALLVNDEANVHLKRINTNSNRIGNRNNNS
KFVIGGVDNPAHVIRFTDDGTKFNFTKQTQGEIVNDFILDAPILPKDLHPDWTNLYIQ
RKILPNDVNTAVVPWPVGRVSGTNADDGMFDFGNGQITNTDPIAQTKTTDNQNPSTF
NSGAMPGANNRYDSQLNVKHRIKTSPQLDEKFVYPEWTGSEENKNITRLATGSLPSNE
RYWIILDIPGTPQVTLKEDSVNVFSRLYLNVSNLSFIGDSIYIFGTSELPSLWYYSFP
TRLSDLTALNQVKTDDIEASSTDNGTTNGTTTGVTSSGSTGAGSAGVTNTNTSQTVSN
PTLNTVRSFGIDSKPTSANKIDETNWADPNVIEARIYABYRLGIQNEIPITNAGNFIR
NTIGGVGFSTGSRVVLRASYNGDQRPTGNFQPFLYVFGYLGYGYQQTRTGTFWYGYKL
LNNSPVDLDAARVGTETNQFRRTSLTYPVMGGYLTEEGARSFSNTPYIRAQGDTPES
RSIFOSGYSDNTYEYIQSVLGPDGIRNNLNVGVKASSFLNSNRPNPNGLEMIAATTYL
RSQIGLARTSGLPNQQPFGTTHQVISVSPGDQFSSIKNIRTIFPGNQLWYFLFTNENN
KSSVYTLRLADSSNPDASSSFPTSLIDVNEIGVLPLLDNSFYVNAAGNVALFSSN
PGSPGSYTAVNTFNQNLSDIAFEGSGAKYTSDFWGTIQFKPDEVLIQNGFTSQVARNF
VTNQSFLNSLVDFTPANAGTNYRVRVVDPDGNLTNQNLPLKVQIQYLDGKYYDAKLKNN
NLVTFSYNNFAALPSWVPTAIGSTLGILAIMIILGLAIGIPLRAQRKLQDKGFKTTF
KKVDTLTAAVGSVYKKIITQTANVKKPAALGAGKSGDKKPAAAAKPSAPKA
SSPAKPTAPKSGAPTKPTAPKPAAPKTAPKE"

gene   7128..8126
            /gene="crmB"
CDS   7128..8126
            /gene="crmB"
            /note="CrmB; similar to N-terminus of Mycoplasma
            gallisepticum GapA"
            /codon_start=1
            /transl_table=4
            /product="putative cytadherence related protein B"
            /protein_id="AAF25383.1"
            /db_xref="GI:6694737"
            /translation="MLKAHIKRSLILFIILLIAAISLIFIGLFGFKKPNINQVKLNNQ
            SSLITNQPSDQIIRKYSLDSIRWNANANFSSSPTKDGGLLVITTDGRLERIDSPGYL
            KWELNLRNINALATVDEEQNLVYSDNRIVIKPLDAVFINNVSKKVVAERAIQSQEDPTI
            YYLLAVSDQIITNDLKDQLLSFQQLATDLTSQGVVLKVKENPTQISDPVWDTYAMIS
            PTDLTQNYPASWKTNTLSLVGDNGSFRKADHPSWFNNDSNFISLPWLQYITNLANLYE
            YKDNVFLFGGNGNWNAIRVFVSKQPNTFLLGCLESGFIKILVLDQLSQVILMLTY"

ORIGIN

Query Match   98.6%; Score 3143.4; DB 1; Length 8354;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3159; Conservative 2; Mismatches 28; Indels 0; Gaps 0;

Qy   1 ATGAAATATTTCTAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3778 ATGAAATATTTCTAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGA 3837

Qy   61 GCTCTTGGTTCTCGCAAGCTTTGGCTTAAGCAATCAGATAAGAGTAACGATAACACGGCAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3838 GCTCTTGGTTCTCGCAAGCTTTGGCTTAAGCAATCAGATAAGAGTAACGATAACACGGCAA 3897

Qy   121 TTAGTTAATCAAGCAAGACGCTAGATGCTAATTCTGTTGAGACTTCTGCAGGTCTTGGACAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3898 TTAGTTAATCAAGCAAGACGCTAGATGCTAATTCTGTTGAGACTTCTGCAGGTCTTGGACAA 3957

Qy   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3958 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 4017

Qy   241 AATGGAACAATTATCATCAAATTACAAATTAGAGTAGTTTACTAAACAAATTGGTTACTTGAT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4018 AATGGAACAATTATCATCAAATTACAAATTAGAGTAGTTTACTAAACAAATTGGTTACTTGAT 4077

Qy   301 GATGGATTGTTGGTGGATGCTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4078 GATGGATTGTTGGTGGATGCTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 4137

Qy   361 AGATTTGATCAAGAACAACAAGACGCATATTATGCTCTCGTTGTTGGTTAATGATGAAGCTAAC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4138 AGATTTGATCAAGAACAACAAGACGCATATTATGCTCTCGTTGTTGGTTAATGATGAAGCTAAC 4197

Qy   421 GTTCATTTAAAAAGAATTAATACTACTAACTGCAATGGTAATAGAGAAACAACAATTCT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4198 GTTCATTTAAAAAGAATTAATACTACTAACTGCAATGGTAATAGAGAAACAACAATTCT 4257

Qy   481 AAGTTTGTAATTGGTGGTGTTGATGATAATCCAGCTCACGTAACTGAATTACTGATGATGGG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4258 AAGTTTGTAATTGGTGGTGTTGATGATAATCCAGCTCACGTAACTGAATTACTGATGATGGG 4317

Qy   541 ACTAAATTAATTTACAAACCAAACTCAAGTGAAATTGTTAAGACTTCATTTTAGAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4318 ACTAAATTAATTTACAAACCAAACTCAAGTGAAATTGTTAAGACTTCATTTTAGAT 4377

Qy   601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4378 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG 4437

Qy   661 ATCTTACCAAATGACGTCAACACTGCGAGTTGTTCCTGGCCAGTAGGTAGTAGTGGA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4438 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTGACCAGTAGGTAGTAGTGGA 4497

Qy   721 ACAAATGCTGATGATGGGATGTTTGATTGTGTGGGAATGGTCAAATAATAACTAATACAGATCCT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4498 ACAAATGCTGATGATGGGATGTTTGATTGTGTGGGAATGGTCAAATAATAACTAATACAGATCCT 4557

Qy   781 ATTGCTCAAACTATAAACCACTACTGACTTCCTTCAACTTTAATTCAGGAGCA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4558 ATTGCTCAAACTATAAACCACTACTGACTTCCTTCAACTTTAATTCAGGAGCA 4617

Qy   841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAAGTCAAGCATAGAATTAAAACA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4618 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAAGTCAAGCATAGAATTAAAACA 4677

Qy   901 TCTTTCCAATTAGATGAAAAAATTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4678 TCTTTCCAATTAGATGAAAAAATTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAA 4737

Qy   961 AATATTACAAGATTAGTACTGGAAGTTTGCCAAGCAACGACGAAAGATATTGGATTCTTGAC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4738 AATATTACAAGATTAGTACTGGAAGTTTGCCAAGCAACGACGAAAGATATTGGATTCTTGAC 4797

Qy   1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTGAACGTATTTTCAAGACTA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4798 ATACCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTGAACGTATTTTCAAGACTA 4857

Qy   1081 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTGATATTATATTTTGTACC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4858 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTGATATTATATTTTGTACT 4917

Qy   1141 TCTGAATTACCATCATTATGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4918 TCTGAATTACCATCATTATGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT 4977

Qy   1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCTTCAAGCACTGATACAACAACA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4978 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCTTCAAGCACTGATACAACAACA 5037

Qy   1261 AACGGAACAACAACAACAACACTTCTAGTGGTTCAACAGTGCTGGAACAGGAAAAT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5038 AACGGAACAACAACAACAACACTTCTAGTGGTTCAACAGTGCTGGAACAGGAAAAT 5097

Qy   1321 ACTACTAACAACTTCTCAAACACTTCTCTAATCCTACTTAAATACTTAAATACTCGTAGTTTTGGA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5098 ACTACTAACAACTTCTCAAACACTTCTCTAATCCTACTTAAATACTTAAATACTCGTAGTTTTGGA 5157

Qy   1381 ATTGATAGTAAACAACTTCTGCAAACAAAATAGTGAAACTAATTGGGCAGATCCTAAC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5158 ATTGATAGTAAACAACTTCTGCAAACAAAATAGTGAAACTAATTGAGCAGATCCTAAT 5217

Qy   1441 GTTATTGAAGCAAGAATATATGCTGAATACAGTATTAGCTATTCAAAATGAAATTCCAATA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5218 GTTATTGAAGCAAGAATATATGCTGAATACAGTATTAGCTATTCAAAATGAAATTCCAATA 5277

Qy   1501 ACTAATGCAGGAAACTTATCCGAAACACAATTGGTGGTGTTTACTTCAACAGGT 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5278 ACTAATGCAGGAAACTTATCCGAAACACAATTGGTGGTGTTTACTTCAACAGGT 5337

Qy   1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAAACCGTAATCCAACTGGAAACTTCCAA 1620

```
Db   5338 TCAAGAGTAGTTTCTTATAACCGTGATCAACGTCCAACTGAAACTTCCAA 5397

Qy   1621 CCTTTCTTATACGTACGTATTGGTTATTTAGGAGTATACCAACAGAACACAGAACTTCTGG 1680
          ||||||||||||||||| ||||||  ||||||||||||||||||||||||||||||||||
Db   5398 CCTTTCTTATACGTACGTATTTGGTTATTTTAGGATACCAACAACTAGAACACAGAACTTCTGA 5457

Qy   1681 TACGGAACATATAAGCTTAAACAACACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA 1740
          |||||||| |||||   ||||||||||||||||||||||||||||| ||||||||||||
Db   5458 TACGGAACATATAAACATTATTAAAACACAGCCCTTACGACGTATTAGATGTGCAAGAGTA 5517

Qy   1741 GGTACTGAAACCAATCAATTAGAGAACTTCATTAACACTTCATTAACATCCCTGTTATGGGTTGGATAT 1800
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Db   5518 GGTATTCTTGCAATTATGATCATCTTAGGAGTATCGGATATCCTTAAGAGCTCAA 5577

Qy   1801 CTTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTTCCATATATAGAGCACAAGGTGAC 1860
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   5578 CTTAACTGAAGAGAGGTGCTAGAAGTTTCTCTAATACTTCCATATATAAGAGCACAAGGTGAC 5637

Qy   1861 ACACCAGAAAGCCGAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT 1920
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   5638 ACACCAGAAAGCCGAGCATCTTCCAATCTGCTAYTCTGATAATACTTATGAGTACATT 5697

Qy   1921 CAATCAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAGCATCA 1980
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   5698 CAATCAGTTTTAGGATTTGATGATTAGAAATTAGAAATAACTTAAATGTTGGGGTTAAGCATCA 5757

Qy   1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5758 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA 5817

Qy   2041 TACTTAAGATCACACAAATTGGATTAGCTAGACATCTGGATTACCAAACCAACCATTC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5818 TACTTAAGATCACACAAATTGGATTAGCTAGACATCTGGATTACCAAACCAACCATTC 5877

Qy   2101 GGAACAACTCACCAAGTATTCAGTATCACCTGGTATCAGTTCTCATCAATTAAGAAT 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5878 GGAACAACTCACCAAGTATTCAGTATCACCTGGTATCAGTTCTCATCAATTAAGAAT 5937

Qy   2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCCTATTCACAAAATGAAAATAAT 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5938 ATTAGAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCCTATTCACAAAATGAAAATAAT 5997

Qy   2221 AAATCTAGTGTTTATACACTAGTTGTTATATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5998 AAATCTAGTGTTTATACATTAGTTGTTATATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 6057

Qy   2281 TTCAGTCCAACAGTTAATTGACGTTAATGACGTAAATGGTGTAAATCTTACCTTTATTAGAGAC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6058 TTCAGTCCAACAAGTTAATTGACGTTAATGACGTAAATGGTGTAAATCTTACCTTTATTAGAGAC 6117

Qy   2341 AATTCATTCTATACAGTAGTCTGCTGCGTGGTAAATGCTGCATTGTTGCATTGTTCTCATCAATGTTGTTCTCATCAAACCCTGGT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6118 AATTCATTCTATACAGTAGTCTGCTGCGTGGTAAATGCTGCATTGTTGCATTGTTCTCATCAAACCCTGGT 6177

Qy   2401 TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6178 TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 6237

Qy   2461 GAAGGTTCTGGTGCTAAGTAGTATACATCTGATTTCTGGGAACAATCCAAACCCGAT 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6238 GAAGGTTCTGGTGCTAAATATACATCTGATTTCTGGGAACAATCCAAACCCGAT 6297

Qy   2521 GAGTACTTAATTCAAAATGGGGTTCACTGAGGTCAAGTGGCTAGAAACTTCGTTACGAACCAA 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6298 GAGTACTTAATTCAAAATGGGGTTCACTGAGGTCAAGTGGCTAGAAACTTCGTTACGAACCAA 6357

Qy   2581 AGCTTCTTAAACAGTTTGACTTCAGTTGACTTCACTCCTGCTAATGCTCTGCTACTAACTACCGTGTA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6358 AGCTTCTTAAACAGTTTGACTTCAGTTGACTTCACTCCTGCTAATGCTCTGCTACTAACTACCGTGTA 6417

Qy   2641 GTGGTTGATCCTGATGGTAATTAACAAGCTTTGGCTTAACAAACCAAACCTCAAAAGTTCAGATCCAA 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6418 GTGGTTGATCCTGATGGTAATTAACAAACCAAACCTACCTCTAAAAGTTCAGATCCAA 6477
```

```
Db   5397

Qy   2701 TACTTAGATGGTAAGTAATATTATGATGCGTAAATTAAAGAACAATAATTAGTAACATTCTCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6478 TACTTAGATGGTAAGTAATATTATGATGCGTAAATTAAAGAACAATAATTAGTAACATTCTCT 6537

Qy   2761 TATAACAACTTGGCGCTACCTTACCTTCATGGGTAGTGCTACAGCGCAATGGTAGTAGTACATTA 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6538 TATAACAACTTGGCGCTACCTTACCTTCATGAGTAGTGCTACAGCGCAATGGTAGTAGTACATTA 6597

Qy   2821 GGTATTCTTGCAATTATGATCATCTTAGGAGTATCGGTATTCCTTAAGAGCTCAA 2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6598 GGTATTCTTGCAATTATGATCATCTTAGGAGTATCGGTATTCCTTAAGAGCTCAA 6657

Qy   2881 AGAAAATTACAAGACAAAGGGTCAAAACAACATTCAAAAAAGTGATACCTGACTGCT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6658 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTGATACCTTGACTGCT 6717

Qy   2941 GCTGTTGGTTCAGTTGCTGGTGCTGGTAAATCTGGTATAACGAACCTGGTGCTGCTGCTAAACCTGCT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6718 GCTGTTGGTTCAGTTGCTGGTGCTGGTAAATCTGGTGATAAGAACCTGGTGCTGCTGCTAAACCT 6777

Qy   3001 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTCTAAACCTGCT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6778 GCTGCTTAGGTGCTGGTAAATCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTCTAAACCTGCT 6837

Qy   3061 GCTCCAGCTAAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACAACTCGGGCCTAAA 3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6838 GCTCCAGCTAAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACAACTCGGGCCTAAA 6897

Qy   3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6898 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 6957

Qy   3181 AAAGAATAA 3189
          |||||||||
Db   6958 AAAGAATAA 6966
```

```
Query Match             98.6%;  Score 3143.4;  DB 6;  Length 10651;
Best Local Similarity   99.1%;  Pred. No. 0;
Matches  3159;  Conservative    28;  Mismatches     2;  Indels     0;  Gaps     0;

Qy      1 ATGAATATTTCTAAAAAACTAAAAACTTAAAGTTATACATTGATTAGGTGGATTAGCTGTATTTGGA 60
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      6 ATGAATATTTCTAAAAAACTAAAAACTTAAAAGTTATACATTGATTAGGTGGATTAGCTGTATTTGGA 65

Qy     61 GCTCTTGGTTCTGCAAGCTTTGGCTTAAGCAATCAGATCAATCAGATTAACGAATAACACGCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 GCTCTTGGTTCTGCAAGCTTTGGCTTAAGCAATCAGATCAATCAGATTAACGAATAACACGCAA 125
```

```
Qy   121  TTAGTTAATCAAGCAAGAACGCTAGATGCTAATCTCGTTAGACTTGCAGTCTTGGACAA  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126  TTAGTTAATCAAGCAAGAACGCTAGATGCTAATCTCGTTAGACTTGCAGTCTTGGACAA  185

Qy   181  AATGGTTCGTCGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186  AATGGTTCGTCGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  245

Qy   241  AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246  AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  305

Qy   301  GATGATTGTGGTGGATACAAAGTAAAACAAAGTAGTTCAGATTACACAACTAGCAGAAAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306  GATGATTGTGGTGGATACAAAGTAAAACAAAGTAGTTCAGATTACACAACTAGCAGAAAT  365

Qy   361  AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCGTTGGTTAATGATGAAGCTAAC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366  AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCGTTGGTTAATGATGAAGCTAAC  425

Qy   421  GTTCATTAAAAAGAATTAATACTAACTCAAATAGGTAATAGAAAACAACAATCT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426  GTTCATTAAAAAGAATTAATACTAACTCAAATAGGTAATAGAAAACAACAATCT  485

Qy   481  AAGTTGTAATTGGTGGTGTTGATATCCAGCTCACGTAATTGATTTACTGATGATGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486  AAGTTGTAATTGGTGGTGTTGATATCCAGCTCACGTAATTGATTTACTGATGATGG  545

Qy   541  ACTAAATTTAATTTTACAAACCAAGTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546  ACTAAATTTAATTTTACAAACCAAGTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  605

Qy   601  GCGCCAATCTTCACCCTAAAGATTACACCCAGATGGTATAACTTATATCATTCAAAGAAAG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606  GCGCCAATCTTCACCCTAAAGATTACACCCAGATGGTATAACTTATATCATTCAAAGAAAG  665

Qy   661  ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666  ATCTTACCAAATGACGTAAACACTGCAGTTGTTCCTTGACCAGTAGGTAGTGA  725

Qy   721  ACAAATGCTGATGATGGGATGTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCT  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726  ACAAATGCTGATGATGGGATGTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCT  785

Qy   781  ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786  ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA  845

Qy   841  ATGCCTGGTGCAAACAATAGTAGCTACTGGAAGTTTGCCAAGCAACGCAAAGTATTGGATT  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846  ATGCCTGGTGCAAACAATAGTAGCTACTGGAAGTTTGCCAAGCAACGCAAAGTATTGAATT  905

Qy   901  TCTTTCCAAATTAGATGAAAAATTTGTTTATCCGAATGACTCGGTTCTCGAAAGAATAAA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906  TCTTTCCAAATTAGATGAAAAATTTGTTTATCCGAATGACTCGGTTCTCGAAAGAATAAA  965

Qy   961  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGAATTCTTGAC  1025

Qy   1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 ATACCAGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTTCAAGACTA  1085

Qy   1081 TACTTAAACTCAGTTAATTCTTATCATTCATTCATTGGTGATAGTATTTTTATATTTTGGTACC  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 TACTTAAACTCAGTTAATTCTTATCATTCATTCATTGGTGATAGTATTTTTATATTTTTGGTACT  1145

Qy   1141 TCTGAATTACCATCATTATGGTACTATTCATTCATTCCCAACTAGATAATCTGATCTAACCGCT  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1146 TCTGAATTACCATCATTATGATACTATTCATTCATTCCCAACTAGATAATCTGATCTAACCGCT  1205
```

```
Qy   1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1206 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA  1265

Qy   1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1266 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1325

Qy   1321 ACTACTAACACTTCTCAAACAGTTCTATCCTATCTTTAAATACTTATCGTAGTTTTGGA  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1326 ACTACTAACACTTCTCAAACAGTTCTATCCTATCTTTAAATACTTATCGTAGTTTTGGA  1385

Qy   1381 AATTGATAGTAACCAACTTCTGCAAACAAATAGATGAAACTAATTGGGCAGATCCTAAC  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1386 ATTTGATAGTAAACCAACTTCTGCAAACAAATAGATGAAACTAATTGAGCAGATCCTAAT  1445

Qy   1441 GTTATTGAAGCAGAATATATGCTGATAACAGATTAGGTATTCAAAATGAAATTCCAATA  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1446 GTTATTGAAGCAAGAATATATGCTGATAACAGATTAGTTATTCAAAATGAAATTCCAATA  1505

Qy   1501 ACTAATGCAGGAAACTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1506 ACTAATGCAGGAAACTTTAAGAGCTTCTTATAACGGTGGTTTACTTCAACAGGT  1565

Qy   1561 TCAAGAGTAGTTTAAGAGCTTCTGTATTTGGTTATTGGAATAGATACTCAACTGGAAACTTCCAA  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1566 TCAAGAGTAGTTTAAGAGCTTCTGTATTTGGTTATTGGTTAGGATCAACAACAACTTCCAA  1625

Qy   1621 CCTTTCTTATACGGTATTTGGTTATTTGGTTAATGACTTCATTTTAGAT  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1626 CCTTTCTTATACGGTATTTGGTTATTTGGTTAATGACTTCATTTTCTGA  1685

Qy   1681 TACGGAACATATAAGCTTTAAACAACAGCCCTACGCGTATTAGATTCTCAAGAGTA  1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1686 TACGGAACATATAAACATTAAACAACAGCCCTTACGCGTATTAGATGCTCAAGAGTA  1745

Qy   1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATAT  1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1746 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATAY  1805

Qy   1801 CTAACTGAAGAAGGTGCTAGAAGTTCTCTAATAACTAATACTCCATATAATAAGAGAGCACAAGGTGAC  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1806 CTAACTGAAGAAGGTGCTAGAAGTTCTCTAATAACTAATACTCCATATAATAAGAGAGGTCAAGGTGAC  1865

Qy   1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTAATCTTGATAATATACTTATGAGTACATT  1920
          |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   1866 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTAYTCTGATAATATACTTATGAGTACATT  1925

Qy   1921 CAATCAGTTTAGGATTTGATGGAATTGAAATAACTTAAATGTTGGGGTTAAAGCATCA  1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1926 CAATCAGTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA  1985

Qy   1981 AGTTTCTTAAACTCAATTAGACCAAATCCAACAAATCTGGAAAATGAATGCTGCTCAACAACA  2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1986 AGCTTCTTAAACTCAATTAGACCAAATCAACAAATCTGGAAATGATTGCTGCTCAACAACA  2045

Qy   2041 TACTTAAGATCACAAAATTGGATTAGCTAGTAGCTGATTCAGATTACCAACCAACAACCATTC  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2046 TACTTAAGATCACAAAATTGGATTAGCTAGTAGCTGATTCAGATTACCAACCAACAACCATTC  2105

Qy   2101 GGAACAACTCACCAAGTTATTCAGTATCACAGTTCTCATTCTTATTCAAATGACAATTAAGAAT  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2106 GGAACAACTCACCAAGTTATTCAGTATCACAGTTCTCATTCTTATTCAAATGACAATTAAGAAT  2165

Qy   2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAAATAAT  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2166 ATTAGAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCACAAATGAAAAATAAT  2225

Qy   2221 AAATCTAGTGTTTATACATTAAGATTAGCTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2226 AAATCTAGTGTTTATACATTAAGATTAGCTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2285

Qy   2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGAC  2340
```

```
REFERENCE   1
AUTHORS     Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki,Y.
TITLE       Modified dna molecule, recombinant containing the same thing, and
            uses thereof
JOURNAL     Patent: EP 1275716-A 79 15-JAN-2003;
            Zeon Corporation (JP)
FEATURES    Location/Qualifiers
source      1..3129
            /organism="Mycoplasma gallisepticum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:2096"
            /note="Modified mgc3 gene"
ORIGIN

Query Match       95.9%;   Score 3057.8;  DB 6;  Length 3129;
Best Local Similarity 98.7%;  Pred. No. 0;
Matches 3083;  Conservative   0;  Mismatches  42;  Indels   0;  Gaps   0;
```

```
Qy     65 TTCGTTCTGCAAGCTTTGGCTTAAGCAATCAGAATCAGATAAGAGTAACGATAACACGCAATTAG 124
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      5 TCCGTTCTGCAAGCTTTGGCTTAAGCAATCAGAATCAGATAAGAGTAACGATAACACGCAATTAG  64

Qy    125 TTAATCAAGCAAGAACGCTAGATGCTAATTCGTTGAGTCTTGCAGGTCTTGCAGGTCTTGACAAAATG 184
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 TTAATCAAGCAAGAACGCTAGATGCTAATTCGTTGAGTCTTGCAGGTCTTGCAGGTCTTGACAACAGG 124

Qy    185 GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCAGTCTAATG 244
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125 GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCAGTCTCAGG 184

Qy    245 GAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTTAAGTGATG 304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 GAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTTAAGTGATG 244

Qy    305 ATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACCAACTAGCAGAAATAGAT 364
          |||| |||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 ATTTTGCTGGATACAAAGTAAAACAAATAGTTTCAGATTACCAACTAGCAGAAATAGAT 304

Qy    365 TTGATCAAGAGAAACAAGAGAGCATATATGCTCTCGTTGGTTAATGATGAAGAAGCTAACGTTC 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305 TTGATCAAGAGAAACAAGAGAGCATATATGCTCTCGTTGGTTAATGATGAAGAAGCTAACGTTC 364

Qy    425 ATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGT 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 ATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACCAAAATTCTAAGT 424

Qy    485 TTGTAATTGGTGGTGTTGATAATCCAGTCACGTAGATTACTGATGATGGGACTA 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425 TTGTAATTGGTGGTGTTGATAATCCAGTCACGTAGATTACTGATGATGGGACTA 484

Qy    545 AATTAATTTACAAACCAAACTCAAGGTGAAATTGTAATGTTAATCAATTTTAGATGCGC 604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    485 AATTCAGTTACACAGCAAACTCAAGGTGAAATTGTAATGTTAATCAATTTTAGATGCGC 544

Qy    605 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATATCAAAGAAAGATCT 664
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    545 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATATCAAAGAAAGATCT 604

Qy    665 TACCAAAATGACGTCAACAACTGCAGCTCCTGGCCAGTAGGTTAGTGGGACAA 724
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    605 TACCAAAATGACGTCAACAACTGCAGCTCCTGGCCAGTAGGTTAGTGGGACAA 664

Qy    725 ATGCTGATGATGGGATGTTTGATTGTGGCGAATGGTCAAATAACTAATAACAGATCCTATTG 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    665 ATGCTGATGATGGGATGTTTGATTGTGGCGAATGGTCAAATAACTAATAACAGATCCTATTG 724

Db    785 CTCAAACTAAAACCACTACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGC 844

Db    725 CTCAAACTAAAACCACTACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGC 784

Qy    845 CTGTTGCAAACAATGACATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2286 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGAC 2345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2341 AATTCATTCATACAGTAAATGCTGCTGGTAAATGCATTGTTCTCATCAAACCCTGT 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2346 AATTCATTCATACAGTAAATGCTGCTGGTAAATGCATTGTTCTCATCAAACCCTGT 2405

Qy   2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2406 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2465

Qy   2461 GAAGGTTCTGCTGCTAAGTATACATCTGATTTCTGGGGACAATCCAATTCAAACCCGAT 2520
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   2466 GAAGGTTCTGCTGCTAAGTATACATCTGATTTCTGGGGACAATCCAATTCAAACCCGAT 2525

Qy   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2526 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2585

Qy   2581 AGCTTCCTTAAACAGTTGACTTAGTTAGTGCTAATGCTCGTGATCTGCTACTACCGTGTA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2586 AGCTTCCTTAAACAGTTGACTTAGTTAGTGCTAATGCTCGTGATCTGCTACTACCGTGTA 2645

Qy   2641 GTGGTTGATCCTGATGGGTAATTAACAAACCAAAACCTCAAAAGTTCAGATCCA 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2646 GTGGTTGATCCTGATGGGTAATTAACAAACCAAAACCTCAAAAGTTCAGATCCA 2705

Qy   2701 TACTTAGATGTGTAAGTAATGATTATGATGATGATGCCTAAATAAAGAACAATATTAGTAACATTCTCT 2760
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   2706 TACTTAGATGTGGTTCAGTTGATGATGATGCCTAAATAAGAACAATAATTAGTAACATTCTCT 2765

Qy   2761 TATAACAACTTGGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2766 TATAACAACTTGCTGCTGCTTTACCTTCATGAGTAGTGCCTACAGCAATTGGTAGTACATTA 2825

Qy   2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCGGTATCCGTTAAGAGCTCAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2826 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCGGTATCCGTTAAGAGCTCAA 2885

Qy   2881 AGAAAATTACAAGACAAAGGGTTCAAAACATTCAAAAAGTTGATACCTTGACTGCT 2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2886 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT 2945

Qy   2941 GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACGTCTAACGTTAAGAAAACCT 3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2946 GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACGTCTAACGTTAAGAAAAACCT 3005

Qy   3001 GCTGCTTAGTGCTCGGTAAATCTGGTGATGATAAGAAACCTGCTGCTGTCTAAACCTGCT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3006 GCTGCTTAGTGCTCGGTAAATCTGGTGATGATAGAAACCTGCTGCTGTCTAAACCTGCT 3065

Qy   3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3066 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA 3125

Qy   3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3126 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3185

Qy   3181 AAAGAATAA 3189
          |||||||||
Db   3186 AAAGAATAA 3194
```

```
RESULT 6
AX665242
LOCUS       AX665242          3129 bp    DNA        linear    PAT 26-MAR-2003
DEFINITION  Sequence 79 from Patent EP1275716.
ACCESSION   AX665242
VERSION     AX665242.1   GI:29290367
KEYWORDS
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
```

Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

```
Db   785 CTGGTGCAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT  844

Qy   905 TCCAATTAGTAGACAAAAATTGTTTATCCAGAATGGACTGGTTCTGAGAGAATAAAAATA  964
         |||||||||||| ||  || |||||||||||||||||||||||||||||||||| ||||
Db   845 TCCAATTAGTAGATGAATGAAAAAATTGTTTATCCAGAATGGACTGGTTCTGAGAGAATAAACAGA  904

Qy   965 TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGAATATTGGATTCTTGACATAC 1024
         |||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||||
Db   905 TTACAAGATTAGCTACTGGAAGTTGCCCAGCAACGAACGAAGAATATTGGATTCTTGACATAC  964

Qy  1025 CCGGGACTCCACAAGTTACTTAAAAAGAAGAGATTCAGTTAAACGTATTTTCAAGACTATACT 1084
         |||||||||||||||||||||| |||||||||||||||||||   |||||||||||||||||
Db   965 CCGGGACTCCACAAGTTACTTTAAAAAGAAGAGATTCAGTTAACGTATTTTCAAGACTATACT 1024

Qy  1085 TAAACTCAGTAATTCTTTATCATTCATTGGTGTAGATGTATTTTATATTTTTTGGTACCTCTG 1144
         |||||||||||||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db  1025 TAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTAGTATTTTATATTTTTTGGTACCTCTG 1084

Qy  1145 AATTACCATCATTATGGTACTATCATTCCCAACTAGATTATCTGATCTAACCGCTTTGA 1204
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  1085 AATTACCATCATTATGGTACTAATCATTCCCAACTAGATTATCTGATCTAACCGCTTTGA 1144

Qy  1205 ATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACAAACG 1264
         ||||||||||||| |||| |||||||||||||||||||||| |||||||||||||||||||
Db  1145 ATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATATCGGTACAACAACAACAGG 1204

Qy  1265 GAACAACGACAACAAGCTGCGATACATCTAGTGGTTCAACAGGTGCTGGAACAGAAATACTA 1324
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  1205 GAACAACGACAACAACTGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGAGACTA 1264

Qy  1325 CTAACACTTCTCAAACAGTTCTCAATCCTAAAAAAATACTTATCGTAATCGTAGTAGTTTTGGAATTG 1384
         ||||||||||||||||||||||||||||||| ||||||| ||| |||||| ||||||||||||||
Db  1265 CTCGACTTCTCAAACAGTTTCTCAAATCCTAAAAATCCTAAAATCTAATCTATCGTAGTTTTTGGAATTG 1324

Qy  1385 ATAGTAAACCAACTTCTGCAAACAAAATAGAAGCAATAATTAATTGGCAGATCCTTAACGTTA 1444
         ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  1325 ATAGTAAACCAACTTCTGCAAACAAAATAGAATGATGAAATAATTGGCAGATCCTTAACGTTA 1384

Qy  1445 TTGAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1504
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  1385 TTGAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1444

Qy  1505 ATGCAGGAAACTTTATCCTGAAAACAACAACAATTGGTGGTGTGTTTACTTCAACAGGTTCAA 1564
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  1445 ATGCAGGAAACTTTATCCTGAAAACAACAACAACAATTGGTGGTGTGTTTACTTCAACAGGTTCAA 1504

Qy  1565 GAGTAGTTTAAGAGCTTCCTCATATACTGTTGGGTTTGGTTCTTCCAACTGGAAACTTCCAACTT 1624
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1505 GAGTAGTTTAAGAGCTTCCTCATATACTGTTGGGTTTGGTTCTTCCAACTGGAAACTTCCAACCTT 1564

Qy  1625 TCTTATACGTATTGGTATTTAGGATACGCACAACAGAAAACAACTAGAACAGGAGAACTTTCTGGTACG 1684
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  1565 TCTTATACGTATTGGTATTTAGGATACGCACAACAGAACGAACAACTAGAACAGGAGAACTTTCTGGTACG 1624

Qy  1685 GAACATATAAGCTTTTAAAACAACAGCCCTATTCATTAACATACCCGTTATGGGGTGGATATCTAA 1744
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  1625 GAACATATAAGCTTTTTTACAGAACAGCCCTATTCATTAACATACCCGTTATGGGGTGGATATCTRA 1684

Qy  1745 CTGAAGAAGGTGCTAGAAGTTCTAGAACTTCTCTAATACTCCATATATAAGAGCACAAGGTGACACAC 1804
         |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  1685 CTGAAAACCAATCAATTTTAGAGAAGTTCTAGAACTTCTCTAATACTCCATATATAAGAGCACAAGGTGACACAC 1744

Qy  1805 CTGAGAAGGTGCTAGAAGTTCTTCTAGAAGTTCCAAATCTGATAATAATACTTATGAGTACATTCAAT 1864
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  1745 CTGAAGAAGGTGCTAGAAGTTCTAGAAGTTCCAAATCTGATAATAATACTTATGAGTACACATTCAAT 1804

Qy  1865 CAGAAAGCCGGAGCATCTTCCAATCGGCCAATTCGGCCATTCTGGATATAATACTTATGATGAGTACATTCAAGCT 1924
         ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  1805 CAGAAAGCCGGAGCATCTTCCAATCGGCCAATTCGGCCATTCTGGATATAATACTTATGATGAGTACATTCAAGCT 1864

Qy  1925 CAGTTTTTAGGATTTGATGGAATTGAAATTGAGAAATAACTTAAAATGTTGGGGTTAAGTGGGGTTAAAGCATCAAGCT 1984
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1865 CAGTTTTTAGGATTTGATGGAATTGAAATTGAGAAATAACTTAAATCGTTGGGGTTAAATGTTGGGGTTAAAGCATCAAGCT 1924
```

```
Qy  1985 TCTTAAACTCAAATAGACCAAATCCAATGATTGCTGCAACAACATACT 2044
         |||||||||||||||||||||||||||||||||| ||||||||| ||||
Db  1925 TCTTAAACTCAAATAGACCAAATCCAAATGATTGCTGCAACAACTACT 1984

Qy  2045 TAAGATCACAAATTGGATTAGCTAGACATCTGGATAGACTAGAACATTCGGAA 2104
         ||||||||||||||||||||||| |||||||||||| |||||||||||||||
Db  1985 TAAGATCACAAATTGGATTAGCTAGACATCTGGATTAGACTAGACTCCATTCCGGAA 2044

Qy  2105 CAACTCACCAGTATTCAGTGATCACCTGGTGATCAGTTCTCATCATCAATTAAGAATATTA 2164
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  2045 CAACTCACCAGTATTCAGTATCACCTGGTGATCAGTTCTCATCATCAATTAAGAATATTA 2104

Qy  2165 GAACAATCTTCCCTGGTAACCAGTATATGGTACTTCTTATTCTTATTCACAAATGAAAATAATAAAT 2224
         ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  2105 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCTTATTCACAAATGAAAATCAGAAAT 2164

Qy  2225 CTAGTGTTTATACATTAAGATTAGTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCATTCA 2284
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  2165 CTAGTGTTTATACATTAAGATTAGTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCATTCA 2224

Qy  2285 GTCCAACAAGTTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGACAATT 2344
         |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  2225 GTCCAACAAGTTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGACAATT 2284

Qy  2345 CATTCTATACAGTGAAATGCTGCTGGTAATGTTGTTCGTCATCCCTGGTTCTC 2404
         |||||||||||||||||||||||||||| |||||||| ||||||||||||||
Db  2285 CATTCTATACAGTGAAATGCTGCTGGTAATGTTGTTCGTCATCCCTGGTTCTC 2344

Qy  2405 CTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCGTGATATTGCTTTTGAAG 2464
         |||||||||||||||| ||||||||||||||| |||||| ||||||||||| |||||||||
Db  2345 CTGGATCATATACTGCTGTAAATACATTTAATCATCAGACAGCGTGTCTGTGATATTGCTTTTGAAG 2404

Qy  2465 GTTCTGGTGCTAAGTAGTATCAATTCGGGGAACAATCCAATTCAAACCCGATGAGT 2524
         ||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  2405 GTTCTGGTGCTAAGTAGTATCATATCGATTTCTGATTTCTGGGGACAATCCAATTCAAACCCGATGAGT 2464

Qy  2525 ACTTAATTCAAAATGGGTTCACTAGTCAAGTGCTAGAAACTTCGTTACAACCAAAGCT 2584
         ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  2465 ACTTAATTCAAAATGGGTTCACTAGTCAAGTGCTAGAAACTTCGTTACACAGCAAAGCT 2524

Qy  2585 TCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGTACTAACTACCGGTAGTGG 2644
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2525 TCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGTACTAACTACCGGTAGTGG 2584

Qy  2645 TTGATCCTGATGGTAATTTAACAACACCAAAAAACCTACCTCTAAAAGTTCAGATCCAAATACT 2704
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2585 TTGATCCTGATGGTCAGCTGACAAACAAAAACCTACCTCTAAAAAGTTCAGATCCAAATACT 2644

Qy  2705 TAGATGGTAAGTATTATGATGCTAAATTAGTAAGAACAATAATTAGTAACATTCTCTATTA 2764
         |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db  2645 TAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCTTATA 2704

Qy  2765 ACAACTTTGGCGCCTTACCTTACTTCATGGGTAGTACAATTGGTAGTACATTAGGTA 2824
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2705 ACAACTTTGGCGCCTTACCTTACTTCATGGGTAGTACAATTGGTAGTACATTAGGTA 2764

Qy  2825 TTCTTGCAATTATGATCATCTTAGGATTAGGGGATGTGGATATCTAA 2884
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2765 TTCTTGCAATTATGATCATCTTAGGATTAGGGGATGTGGATATCTRA 2824

Qy  2885 AATTACAAGACAAAGGGTTCAAAACAACACAATTCAAAAAAGTTGATACCTTGACTGCTGTG 2944
         |||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db  2825 AATTACAAGACAAAGGGGTTCAAAACAACACAATTCAAAAAAGTTGATACCTTGACTGCTGTG 2884

Qy  2945 TTGGTTTCAGTTTACAAGAAGATTATTACCCAAACCGCTAAGTTAAGAAAAACCTGCTG 3004
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  2885 TTGGTTTCAGTTTACAAGAAGATTATTACCCAAACCCAAACTGCTAAGTTAAGAAAAACCTGCTG 2944

Qy  3005 CTTTAGGTGCTGGTGAAATCTGGTGATAAGAAATCTGCTGCTCTAAACCTGCTCGCTC 3064
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2945 CTTTAGGTGCTGGTGAAATCTGGTGATAAGAAATCTGCTGCTCTAAACCTGCTCGCTC 3004
```

```
Matches 1145;   Conservative   4;   Mismatches   81;   Indels   6;   Gaps   4;

Qy     1 ATGAATATTTCTAAAAAACTAAAAGTATACATTGATAGGTGATTAGCTGTATTTGGA    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5910 ATGAATATTTCTAAAAAACTAAAAGTATACATTGATAGGTGATTAGCTGTATTTGGA  5969

Qy    61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACCGCAA   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5970 ACTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACCGCAA  6029

Qy   121 TTAGTAATCAAGCAAGAACGCTAGATGATGCTAATTCTGTTAGACTTGCGAGGTCTTGGACAA   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6030 TTAGTAATCAAGCAAGAACGCTAGATGATGCTAATTCTGTTAGACTTGCGAGGTCTTGGACAA  6089

Qy   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6090 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  6149

Qy   241 AATGGACAATTATCAAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGT   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6150 AATGGACAATTATCAAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGT  6209

Qy   301 GATGATTGGTGGTGGATACAAGAAAACAAAATAGTTTCAGATTACCAACTAGCAGAAAT   360
           ||||||||||    |||||||   ||||||||||||||||||||||||||||||||||
Db  6210 GATGATTTGCTGGATACAAGTAAAACAAAATAGTTTCAGATTACCAACTAGCAGAAAT  6269

Qy   361 AGATTTGATCAAGACAAACAAGAGCATATATGCTCTGTTGGTAATGATGAAGCTAAC   420
           ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  6270 AGATTGATCAAGACAAACAAGAGCATATTATGCTCTGTTGGTAATGATGAAGCTAAC  6329

Qy   421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6330 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT  6389

Qy   481 AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTACTGATGATGGG   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6390 AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTACTGATGATGGG  6449

Qy   541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAA-TGACTTCATTTTAGA   599
           ||||||||||||||||||||   ||||Y|||||||||||||| ||||||| ||||||||
Db  6450 ACTAAATTAATTTACAAAGCAAACYCAAGGTGAAATTGTTAACCGACTTTATTTTAGA  6509

Qy   600 TGGGCCAATCTTCTACCTAAAGATTTACACCCCAGATTGTATAACATTCAAAGAAA   659
           ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db  6510 TGGGCCAATCTTMCCTAAAGATTTACACCCCAGATTGTATAACATTCATTCAAAGAAA  6569

Qy   660 GATCTTACCAAATGACGTCAAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTTGG   719
           |||||||||||||||||| |||||||||||||||||||||| |  ||||||||||||
Db  6570 GATCTTACCAAATGACGTAAACACTGCAGTTGTTCCTTGACCCGTAGGTAGTTGG  6629

Qy   720 AACAAATGCT-GATGATGGGGATGTTGATTGTGGGAATGGTCAAATAATAACTACAGATC   778
           ||  | ||| |||  | |  ||| ||||||||   | ||| || |  |   |  ||||||
Db  6630 TAATAGCGCTACAGGDTGGGACATTTGATTGTTATCCGGGAACGGMATGCCTTCTGCTAATGATC  6689

Qy   779 CTATTGCTCAAACTCAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAG   838
           |  ||||||||||| |  || | |  |  |  ||||||    ||||  ||  |||||||||
Db  6690 CAATAGCAACGTCTAAAGATTTACACCCCAGGTTCTCAAAAT---AMCCCCCCAAGTTCAATTCAAGCG  6746

Qy   839 CAATGCCTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAA   898
           || |  ||||  ||| |||||||| |||||| |  ||||| | |||  | ||||||||||
Db  6747 CTCAACCTGGACAAATAATAGATATGATTCTGAAGTTAATGTAAAACACAGAATTAAAA  6806

Qy   899 CATCTTTCCAATTAGATGAAAAATTTGTTATCCAGATGGTTCTGAAGAGATA   958
           |||| ||||||||||   |||||| ||||||||||||||  |||||||||| |||
Db  6807 CATCGTTCCAATTAGACCGAAAAAATTTGTTATCCAGATGAACTGGTTCTGAAGAGAATA  6866

Qy   959 AAAATATTACAAGATTAGCTACTGGAAGTTTGCCAGCAACAAACAAGATTAGTTCATTCTTG  1018
           |||||||||||||||||||||||||||||||||| | |||||||    |||  || |||||||
Db  6867 AAAATATTACAAGATTAGCTACTGGAAGTTTGCCAAACAACAAGAGGAGATTGAATTCTTG  6926

Qy  1019 ACATACCCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGAC  1078
           |||||||  ||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  6927 ACATACCAGGGACTCCCAGAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGAC  6986
```

```
Qy  3065 CAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAAACCAACTGGGCCTAAATCTG  3124
           |||||||||||||||||||||||||||||||||||||||||||||||  |||||||||||
Db  3005 CAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAAACCAACTGCGCCTAAATCTG  3064

Qy  3125 GTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCGCTCCCAAAG  3184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3065 GTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCGCTCCCAAAG  3124

Qy  3185 AATAA  3189
           |||||
Db  3125 AATAA  3129
```

```
RESULT 7
AF083976
LOCUS       AF083976                7141 bp    DNA     linear   BCT 22-AUG-1998
DEFINITION  Mycoplasma gallisepticum adherence protein gene, complete cds.
ACCESSION   AF083976
VERSION     AF083976.1  GI:3450890
KEYWORDS
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 7141)
  AUTHORS   Goh,M.S., Gorton,T.S., Forsyth,M.H., Troy,K.E. and Geary,S.J.
  TITLE     Molecular and Biochemical Analysis of a 105 kDa Mycoplasma
            gallisepticum cytadhesin
  JOURNAL   Microbiology (1998) In press
REFERENCE   2  (bases 1 to 7141)
  AUTHORS   Goh,M.S. and Geary,S.J.
  TITLE     Mycoplasma gallisepticum adherence protein gene (gapA) and flanking
            regions
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 7141)
  AUTHORS   Goh,M.S. and Geary,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-1998) Pathobiology, University of Connecticut, 61
            North Eagleville Road, Storrs, CT 06269, USA
FEATURES             Location/Qualifiers
     source          1..7141
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="S6"
                     /db_xref="taxon:2096"
     CDS             2994..5888
                     /note="GapA; Mycoplasma pneumoniae P1 homolog; gapA;
                     formerly referred to as MGP1"
                     /codon_start=1
                     /transl_table=4
                     /product="adherence protein"
                     /protein_id="AAC32609.1"
                     /db_xref="GI:3450891"
                     /translation="MQRFNKNKRLKIKCLKTTLCLMHLLVILALKVLQLLLVVVQVV
                     VLLEEVLLLPAASSTARLVEEGNSAGMGTMTPTASTSTVIDYNSDQNKIPKPKTLLD
                     SSESSESINGGRTYANINTQKNLQGVIVKVNENLFNSENPFAVENMAFIKPKIWVDNY
                     PSTWTQGSANGKWTNVLQFYKHDNPAVNNRFYRAKYPKRLETQTTPLIDSSFSPY
                     EHPEWYEGNQFVMPWMQYITNLGGLYAXDGMVYLFGGNGTWVNNESALSIGVFRTKPE
                     NRTAEAPGNTKTVGYPYGILLSAISFDATRNGLALAAPALGQDVGYHFVPRLAVGGVS
                     SPRGANGNIPLGSSAITWGTNGGNFLDTKWHSPAVIEDAPTTFITVNSSGVLQNSGSGQ
                     STSTPMPNSNGNESIPYRWTNSYDYNSVRFAALISKPAGGNTKQVESLFFTALKLDTL
                     NSLPNKFTQENNIPFSYAMLDGRQWSLGTRKDSTWLTNTINNPTYNTQQQLASTAAG
                     ENANPRNILNALTTAKGFDRRDIGNVVTYSNNTNKFTYYQVGGAITTWPEVQVNYK
                     TSANITYYNLTRTDFGSTTPATQDANTVSSKLNGAYLSSTGDQGWYNGSIVKKASF
                     TPSSQGYTWQDFKGLTTTASNAVISNWTKAGYSIRPDDDTVFSVSKIPFEKEITAAVN
                     VRSLDSYYVQLNGETSVNTVARVSPDSSALTINPKRITNPLMNRDNVIGQGAFISRND
                     IPSSFFENNIDIVTTEADGTEVLDSKYINSIYRTYPFQNNPHIRLRLLVIDRSRATN
                     DFIKLLPQVLVDGEYVAVPQANSVFVSDQEFTGFDALPGYVLPVAISIPIIIIIALALA
                     LGLGIGIPMSQNRKMLKQGFAISNKKVDLLTIAVGSVFKQIINRTSVTNIKTPQMLQ
                     ANKKDGASSPSKPSAPAAKKPTGPTKPSAPGAKPTAKPKAPATKKIE"
ORIGIN
```

```
Query Match             33.0%;  Score 1052.4;  DB 1;  Length 7141;
Best Local Similarity   92.6%;  Pred. No. 1.2e-167;
```

```
Qy  121 TTAGTTAATCAAGCAAGAAGACGCTAGAGTCTAATTCTCTGTGATAGTAGTATTATATTTTGGACAA  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TTAGTTAATCAAGCAAGAAGACGCTAGAGTCTAATTCTCTGTGATAGTAGTATTATATTTTGGACAA  180

Qy  181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240

Qy  241 AATGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  300
        ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  300

Qy  301 GATGATTGTTGCTGGATACAAAGTAAAACAAATAGTTCAGATTACACAACTAGCAGAAAT  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GATGATTGTTGCTGGATACAAAGTAAAACAAATAGTTCAGATTACACAACTAGCAGAAAT  360

Qy  361 AGATTTGATCAAGACAAACAAGAGCATATTATGTCTCGTTGGTTAATGATGAAGCTAAC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AGATTTGATCAAGACAAACAAGAGCATATTATGTCTCGTTGGTTAATGATGAAGCTAAC  420

Qy  421 GTTCATTTAAAAGAATTAATACTAACTCAAATAGAATGGTAATAGAAACAACAATTCT  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTTCATTTAAAAGAATTAATACTAACTCAAATAGAATGGTAATAGAAACAACAATTCT  480

Qy  481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGTCCACGTAATTAGATTTACTGATGATGGG  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGTCCACGTAATTAGATTTACTGATGATGGG  540

Qy  541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600

Qy  601 GCGCCAATCTTACCTGAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCGCCAATCTTACCTGAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660

Qy  661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGGA  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ATCTTACCAAATGACGTAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGGT  720

Qy  721 ACAAATGCTGATGATGGGATGTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 AATAGGCTACAGATGGGACATTTGATTCCGGGAACGCATGCCTTCTGCTAATGATCCA  780

Qy  781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTCAACTTTAATTCAGGAGCA  840
        ||||||| ||||||||||||||| ||||||||||||||          ||||||||||
Db  781 ATAGCAACGTCTAAAAGCGGTTCGTAAAATGATGAATTCAATTCAATTCAAGCGCT  837

Qy  841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAAACA  900
        |||||||||| ||||| |||||| |||||||||| ||||||||||| |||||| |||||||
Db  838 CAACCTGGGACAAGAAATTATGATGAAGTTAATGTTAAACACAGAATTAAGAATTAAAACA  897

Qy  901 TCTTTCCAATTAGATGAAAAAATTTGTTTATCCCAGAATGAATGACTCGGTTCTGAAGAGAATAAA  960
        ||||| ||||||||| |||||||||||| ||||||||||||||| |||||||||||||||| |||||
Db  898 TCGTTTCAATTAGACGAAAAATTTGTTTATCCAGAATGAATGACTGGGTTCTGAAGAGAATAAA  957

Qy  961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAACAACAACGAAGATATTCGATTCTTGAC  1020
        ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  958 AATATTACAAGATTAGCTACTGGAAGTTTGCCAACAACAACGAAGATATTGAATTCTTGAC  1017

Qy 1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAACGTATTTTCAAGACTA  1080
        ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1018 ATACCAGGGACTCCAGAAGTTTACTTTAAAAGAAGATTCAGTTAACGTAACGTATTTTCAAGACTA  1077

Qy 1081 TACTTAAACTCAGTTAATTCTTTATCATTGGTGATAGTATTATATTTTT  1134
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1078 TACTTAAACTCAGTTAATTCTTTATCATTGGTGATAGTATTATATTTTT  1131
```

RESULT 9
AB033211
LOCUS AB033211

```
Qy 1079 TATACTTAAAACTCAGTTAATTCTTTATCATTGGTGATAGTATTATATTTTTGGGTA  1138
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6987 TATACTTAAAACTCAGTTAATTCTTTATCATTGGTGATAGTATTATATTTTTGGGTA  7046

Qy 1139 CCTCTGAATTACCACCATCATTATGGTACTATTCCCAACTAGATAATCTGATCTAACCG  1198
        | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7047 CTTCTGAATTACCATCATCATTATGATACTAGAGATTCCCAACTAGATAATCTGATCTAACCG  7106

Qy 1199 CTTTGAATCAAGTTAAAACAGATGATATTGAAGCTT  1234
        |||||||||||   |||||||||||||||||||||
Db 7107 CTTCTGAATCAAGTT-AAACAGATGATATTGAAGCTT  7141
```

RESULT 8
AB033210

LOCUS   AB033210   1131 bp   DNA   linear   BCT 07-APR-2000
DEFINITION Mycoplasma gallisepticum gene for 120-kDa membrane protein MGC3, partial cds, strain:S6.
ACCESSION   AB033210
VERSION   AB033210.1 GI:7527370
KEYWORDS   120-kDa membrane protein MGC3.
SOURCE   Mycoplasma gallisepticum
ORGANISM   Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 1131)
AUTHORS   Yoshida,S.
TITLE   Mycoplasma gallisepticum S6-strain gene encoding a 120-kDa membrane protein
JOURNAL   Published Only in DataBase (2000)
REFERENCE   2 (bases 1 to 1131)
AUTHORS   Yoshida,S.
TITLE   Direct Submission
JOURNAL   Submitted (04-OCT-1999) Shigeto Yoshida, Jichi Medical School, Department of Medical Zoology; Yakushiji 3311-1, Minamikawachimachi, Tochigi 329-0498, Japan (E-mail:shigeto@jichi.ac.jp, Tel:81-285-58-7339, Fax:81-285-44-6489)
FEATURES   Location/Qualifiers
     source   1..1131
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="S6"
                     /db_xref="taxon:2096"
     gene   1..1131
                     /gene="mgc3"
     CDS   1..>1131
                     /gene="mgc3"
                     /codon_start=1
                     /transl_table=4
                     /product="120-kDa membrane protein MGC3"
                     /protein_id="BAA94277.1"
                     /db_xref="GI:7527371"
                     /translation="MNISKKLKSYTLIGGLAVFGTLGSSASFGFKQSDKSNDNTCLVNQ
                     ARTLDANSVRLAGLGQNGSLFNTVLRDVDDNFITAANGTIIKLDSFTKPLYGLDLSDD
                     FAGYKVKQIVSDYTTSRNRFDQRQTRAYALLVNDEANVHLKRINTNSNRIGNRNNNS
                     KPVIGVGVDNPAHVIRFTDDGTKFENFTKQTQGEIVNDPILDAPILPKDLHPDWYNLYIQ
                     RKILPNDVNTAVVPWPVGRVSGNSATDGTFDFGNGMPSANDPIATSKSGSGQNNPSSFN
                     SSAQPGTNNRYDSEVNVKGRIKTSFQLDEKFVYPEWTGSSENKNITRLATGSLPNNEG
                     YWILDIPGTPEVTLKEDSVNVFSRLYLNSVNSLSFIGDSIYIF"
ORIGIN

Query Match 31.1%; Score 993.2; DB 1; Length 1131;
Best Local Similarity 92.9%; Pred. No. 1.7e-157;
Matches 1053; Mismatches 78; Conservative 0; Indels 3; Gaps 1;

```
Qy    1 ATGAATATTTCTAAAAAACTAAAAAGTTATACACATTGATAGGTGGATTAGCTGTATTTGGA  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAATATTTCTAAAAAACTAAAAAGTTATACACATTGATAGGTGGATTAGCTGTATTTGGA  60

Qy   61 GCTCTTGGTGGTTCTGCAAGCTTTGGCTTTAAGCAATCGATAAGAGTAACGTAACACGCCAA  120
        | |||||||||||||||||||||||||||||||||||||| ||||||||||||||  |||||
Db   61 ACTCTTGGTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGTAACAACACGCAA  120
```

```
Qy    361  AGATTTGATCAAAGACAACAAGACAAACAAGAGACATATTATGTCTGTTGGTAATGATGAAGCTAAC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGATTTGATCAAGACAAACAACAAGAGACGATATTATGCTTGTTGGTAGTTAGTAATGATGTCTATC  420

Qy    421  GTTCATTTAAAAGAGATTAATACTAACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GTTCATTTAAAAGAGATTAATAATACTAACCAAATAGAATTGGTAATTAAATAATAATTCG  480

Qy    481  AAGTTGTAATTGGTGGTGTTGATAATCCAGTCACGTAATTAGATTTACTGATGATGGG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  AAATTTGTAATCGGCGGTGTTGATAATCCAGTCACGTAATTAGATTTACTGATGATGGG  540

Qy    541  ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGTTAACGACTTTATTTTAGAT  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  ACTAAATTAATTTTACAAAGCAAACTCAAGGTGAAATTGTTAACGACTTTATTTTAGAT  600

Qy    601  GGGCCAATCTTACCTAAAGATTTACACCCAGTTGTATAATCATTCAAAGAAAG  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GGGCCAATCTTACCTAAAGATTTACACCCAGATTGATACAACTTATCATTCAAAGAAAG  660

Qy    661  ATCTTACCAAATGACGTCAACACTGCAGTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  ATCTTACCAAATGACGTAAACACTGCAGTGTTCCTTGACCTGTAGGTAGTGGTAGTGGT  720

Qy    721  ACAAATGCTGATGATGGATGTTTGATTGTGGGAATGGTCAAAATAACTAATACAGATCCT  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACAAATGCTACTGATGGACATTTGATTTGGTGTAATGGTGTAGCTTTAGTACTACAAACCCT  780

Qy    781  ATTGCTCAAACTAAAAACCACTACTGATAAATCAAAATCCTTCAACTTTAATTCAGGAGCA  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  ATTGAACAAACTAAAAACCACTACCACCAATCAATTGTCTATGACTTTGATTCATCAGCA  840

Qy    841  ATGCCTGGTGCAAACAATAGATACGATTCCAATTGAATGTCAAGCATAGAATTAAAACA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  ATGCCCAGTGCAAATATAGATGATGATTCTGAATTAAAATGTCAAACACAGAATTAAAACA  900

Qy    901  TCTTTCCAATTAGATGAAAAATTGTTTATCCAGAATGCAATGGTTCTGAAGAGAATAAA  960
           ||| |||||||||||||||||||||||||||||||||||||||||||
Db    901  TCGTTTGGATTAGACGAAAAATTGTTTATCCGATAATGAATGGTCAAGAGAGAGAATGCA  960

Qy    961  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020
                                    ------CTAGTGGTAACAATTATTGAATACTTAAT  1014
Db    961  AATCTTACAAGATTAACAAATAGA------CTAGTGGTAACAATTATTGAATACTTAAT  1014

Qy   1021  ATACCCGGACTCCACAAGTTCACTTAAAAGAAGAGATTCAGTTAACGTATTTCAAGACTA  1080
           |||||||||||||||||||||| ||||||||||||||||||||||||
Db   1015  ATTCCTGGAACTCCAAATGTACTTACTTTAAAGAAGACACAGTTAACGTATTTTCAAGATTG  1074

Qy   1081  TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTATTTATATTTTT  1134
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1075  TACTTAAACTCAGTTAACTCTCTATCTTTCATCGGTGATAGTAGTATTTATATTTTT  1128
```

```
RESULT 10
MYCMGP
LOCUS        MYCMGP                    8760 bp    DNA      linear   BCT 26-APR-1993
DEFINITION   M.genitalium attachment protein (MgPa) gene, complete cds.
ACCESSION    M31431
VERSION      M31431.1  GI:150157
KEYWORDS     attachment protein.
SOURCE       Mycoplasma genitalium
  ORGANISM   Mycoplasma genitalium
             Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE    1  (bases 1 to 8760)
  AUTHORS    Inamine,J.M., Loechel,S., Collier,A.M., Barile,M.F. and Hu,P.C.
  TITLE      Nucleotide sequence of the MgPa (mgp) operon of Mycoplasma
             genitalium and comparison to the P1 (mpp) operon of Mycoplasma
             pneumoniae
  JOURNAL    Gene 82 (2), 259-267 (1989)
  MEDLINE    90060815
  PUBMED     2583522
COMMENT      Original source text: M.genitalium (strain G-37, ATCC 33530) DNA.
FEATURES             Location/Qualifiers
```

```
DEFINITION   Mycoplasma gallisepticum gene for 120-kDa membrane protein MGC3,
             partial cds, strain:F.
ACCESSION    AB033211
VERSION      AB033211.1  GI:7527372
KEYWORDS     120-kDa membrane protein MGC3.
SOURCE       Mycoplasma gallisepticum
  ORGANISM   Mycoplasma gallisepticum
             Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE    1  (bases 1 to 1128)
  AUTHORS    Yoshida,S.
  TITLE      Mycoplasma gallisepticum F-strain gene encoding a 120-kDa membrane
             protein
  JOURNAL    Published Only in DataBase (2000)
REFERENCE    2  (bases 1 to 1128)
  AUTHORS    Yoshida,S.
  TITLE      Direct Submission
  JOURNAL    Submitted (04-OCT-1999) Shigeto Yoshida, Jichi Medical School,
             Department of Medical Zoology; Yakushiji 3311-1,
             Minamikawachimachi, Tochigi 329-0498, Japan
             (E-mail:shigeto@jichi.ac.jp, Tel:81-285-58-7339,
             Fax:81-285-44-6489)
FEATURES             Location/Qualifiers
     source          1..1128
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="F"
                     /db_xref="taxon:2096"
     gene            1..1128
                     /gene="mgc3"
     CDS             1..>1128
                     /gene="mgc3"
                     /codon_start=1
                     /transl_table=4
                     /product="120-kDa membrane protein MGC3"
                     /protein_id="BAA94278.1"
                     /db_xref="GI:7527373"
                     /translation="MNISKKLKSYTLIGGLAVLGTLGSASFGFKQSDKSNDNTQLVNQ
                     ARTLDANSVKLAGLGQNGSLFNTVLRDVDDNPITAASGTIIKLDSFTKPLYGLDLSDD
                     FAGYKVKQIVSDYTTSRNFDQRQTRAYYALLVNDDAIVHLKRINTNQNRIGNLNNNS
                     KFVIGGVKDNPAHVIRFTDDGTRFKQTQGEIVNDFILDAPILPKDLHPDWYNLYIQ
                     RKILPNDVNTAVVPWPVGRVSGTNATDGTFDRGNGVAFSTNPIEQTKTTTNNQLSMTF
                     DSSAMPSANNRXDSELNVKKHRIKTSFGLDBKFVYPBWTGQBENANLTRLINRPSGGNY
                     WILNIPGTPNVTLKEDTVNVFSRLYLNSVNSLSFIGDSIYIF"
ORIGIN
Query Match      29.0%;   Score 924.8;   DB 1;   Length 1128;
Best Local Similarity   89.2%;   Pred. No. 5.5e-146;
Matches 1011;   Conservative    0;   Mismatches 117;   Indels    6;   Gaps    1;

Qy      1  ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTAGGA  60

Qy     61  GCTCTTGGTCTCGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     61  ACTCCTGGTCTCGCAAGCTTTGGCTTTAAACAATCAGATAAGAGTAACGATAACACGCAA  120

Qy    121  TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA  180
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    121  TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAAACTTGCAGGTCTTGGACAA  180

Qy    181  AATGTTCCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AATGTTCATTGTTCAATACAGTTCTTAGAGATGTTGATGATAACAACTTTATAACAGCAGCT  240

Qy    241  AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATGGTTAGATCTAAGT  300
           | |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    241  AGTGGAACAATTATCAAATTAGATAGTTTTACTAAAGCCATTATGGTTAGATCTAAGC  300

Qy    301  GATGATTGTGGTGGATACAAAGTAAAACAACAAATAGTTTCAGATTACACAACTAGAAAT  360
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    301  GATGATTTTGCTGGATATAAAGTGAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT  360
```

```
source          1..8760
                /organism="Mycoplasma genitalium"
                /mol_type="genomic DNA"
                /db_xref="taxon:2097"
CDS             298..1059
                /note="29 kD protein"
                /codon_start=1
                /transl_table=1
                /protein_id="AAA25419.1"
                /db_xref="GI:150158"
                /translation="MTFTNQIDESFLKEALAIVVDANYKNRIELRELLDKNLFKAVLR
                IDHIPNEDDLNTSFNFVEBSYVACCEQIVEMATVAKWTIPPVAATLLYIGIYTDSNRF
                LYSNYSYRTLYLAAILYKAKADIRIVHDHLNHTSLADLKPKKVYYNHFKTQGQVIYFI
                CTKKIQKRLRMTADQCARVNLLSNIADYKIWLFFIEQANNEIRIELRSNGINVRDIAI
                KYGGGGHNNASGAIITNKKQISDVVSDCVKKIVYN"
CDS             1066..5400
                /note="attachment protein MgPa"
                /codon_start=1
                /transl_table=4
                /protein_id="AAA25420.1"
                /db_xref="GI:150159"
                /translation="MHQPKKRLAKKSWAFLTAALTLGVITGVGGYFLFNQNKQRSSVS
                NFAYOPKQLSVKHQQAVDETLTPWTWNNNNPSSLKITGENPGSFGLVRSQNDNLNISS
                VTKNSSDDNLKYLNAVEKYLDGQQNFAIRRYDNNGRALYDINLAKMENPSTVQRGLNG
                EPIFDPFKGFGLTGNAPTDWNEIKGKVPVFVVQSPHSPNLYFVLLVPKVALEYHNLNN
                QVVKESLEVKATQSFNPTQRLQKDSPVKDSSKQEKLSETTASSMSSGMATSTRAKA
                LKVEVERGSQSDSLLKGDFAKKPLKHKNSSGEVKLEAEKFTEAWKPLLTTDQIAREK
                GMGATVSFYDAPYSENHTAFGLVDHIDPKKMVENYPPSWKTPKWNHHGIWDYNARNL
                LLQTTGFFNPRRHPEWFDEGQAKADNTSPGFKVGDTDHHKDGFKKNSSSPIALPFEAY
                FANIGNMVAIGNSVFIFGGNGHATKMFTTNPLSIGVFRIKYTPHFSKSSVTGWPYAVL
                FGGLINPQTNGLKDLPLGTNRWFEYVPRMAVSGVKWVGNQLVLAGTLTMGDTATVFRL
                KYDQLEKHLANLVAQQGLLREDLQIFPTPYGWANRPDIPVGAWLQDEMGSKFGPHYFLN
                NPDIODNVNNDTVEALISSYKNTDKLKHVLPYRYSGLYAWQLFNWSNKLTNTPLSANF
                VNENSYAPNSLFAAILNEDLLTGLSDKIFYGKENEFAENEADRFNQLLSLNPNNTNW
                ARYLNVVQRFTTGPNLDSSTPDQFLDFLPWIGNGKPFSNSPSPSTSASSTPLPTFSN
                INVGVKSMITCHLNKENTRWVFIPNFSPDIWTGACYRVQSANQKNGIPFEQVKPSNNS
                TPFDPNSDDNKVTPSCGSSKPTTYPALPNSISPTSDWINALTFTNKNNPQRNQLLLRS
                LLGTIPVLINKSGDSNDQFNKDSEQKWDKTETNEGNLPGFGEVNGLYNAALLHTYGFP
                GTNTNSTDPKIGFKADSSSSSSTLVGSGLNWTSQDVGNLVVINDTSFGFQLGGWFIT
                FTDFIRPRTGYLGITLLSSLQDQTIIWADQPWTSFKGSYLDSDGTPKSLWDPTALKSLP
                NSSTTYDTNPTLSPSFQLYQPNKVKRAYQTTNTYNKLIBPVDATSAATNWTSLLKLTT
                KNIKALGKGTASSQGNNNGGVSQTINTITTGNISEGLKEETSIQAETLKKFPDSK
                QNNKSEIGIGDSTFTKMDGKLTGVVSTPLVNLINGQGATSDSDTEKISFKPGNQIDFN
                RLFTLPVTELFPDNTMFVYDQYVPLLLVNLPSGFDQASIRLKVISYSVENQTLGVRLEP
                KDPQTQOFIPVLNASSTGPQTVPQPFNQWADYVLPLIVTVPIVVIILSVTLGLTIGIP
                MHRNKKALQAGFDLSNKKVDVLTKAVGSVFKEIINRTGISNAPKKLKQATPTKPTPKT
                PPKPPVKQ"
CDS             5402..8560
                /note="114 kD protein"
                /codon_start=1
                /transl_table=4
                /protein_id="AAA25421.1"
                /db_xref="GI:150160"
                /translation="MKTMRKQIYKKAYWLLLPFLPLALANTFLVKEDSKNVTAYTPFA
                TPITDSKSDLVSLAQLDSSYQIADQTIHNTNLPVLFKSRDVKVKYESSGSNNISFDST
                SQGEKPSYVVEFTNSTNIGIKWTMVCKYQLDVPNVSSDMNQVLKNLLEOPLTKYTLN
                SQGEKPSYVVEFTNSTNIGIKWTMVCKYQLDVPNVSSDMNQVLKNLLEOPLTKYTLN
                SSLAKEKGKTQREVHLGSGQANQWTSQRNQHDLNNNPSPNASTGFKLTTGNAYRKLSE
                SWPIYEPIDGTKQGKGKDSGGWSSTBENEAKNDAPVSGGGSSSGTFNKYLNTKQALE
                SIGILFDDQTPRNVITQLYYASTSKLAVTNNHIVVMGNSFLPSMWYWVVERSAQENAS
                NKPTWFANTNLDWGEDKQKQFVENQLGYKETTSTNSHNFHSKSFTQPAYLISGIDSVN
                DQIIFSGFKAGSVGYDSSSSSSSSTKDQALAWSTTTSLDSKITGYKDLVTNDTGLN
                GPINGSPSIQDTFSFVVPYSGNHTNNGTTGPIKTAYPVKDQKSTVKINSLINATPLN
                SYGDEGIGVFDALGLNYNFKSNQERLPSRTDQIFVYGIVSPNELRSAKSSADSTGSDT
                KVNWSNTQSRYLPVPYNYNSEGIIDADGFKRDENRGASVTTFSGLKSIAPDGFANSIAN
                FSVGLKAGIDPNPVMSGKKANYGAVYRLNFNPGNDSLLSTTDNNIAPISFS
                FTPFTAAESAVDLTTKEVTVNQESGLWSYIFDSSLKPSHDGKQIPVTDNMGFSVITV
                SRTGIELNQDQATTLDVAPSALAVAVOSGIQSTTQTIGVLPLSEEFSAVIAKDSDQNK
                IDIYKNNNGLFEIDTQLSNSVATNNGGLAPSTENRVDAWGKVEFADNSVLQARNLVD
                KTVDEIINTPEIINSPRFTPAFEDKATLVATKQSDTSLSVSPRIQFLDGNFYDLNS
                TIAGVPLNIGFPSRVFAGFAALPAWIVPVSVGSSVGIIPFILLVLGLGIGIPMYRVRKL
                QDASFVNVFKKVDTLITAVGSVYKKIITQTGVKKAPSALKAANPSVKKPAAFLKPPV
                QPPSKPFGEQKAVEVKSEETKS"
misc_signal     8582..8613
```

```
                    /organism="Mycoplasma genitalium"
                    /mol_type="genomic DNA"
                    /isolate="G37"
                    /db_xref="taxon:2097"
gene     1264..2016
                    /gene="MG186"
CDS      1264..2016
                    /gene="MG186"
                    /note="similar to GB:U00089 SP:P75265 PID:1673668 percent
                    identity: 44.35; identified by sequence similarity;
                    putative"
                    /codon_start=1
                    /transl_table=4
                    /product="liproprotein, putative"
                    /protein_id="AAC71405.1"
                    /db_xref="GI:3844783"
                    /translation="MKGFLKPNFSLGALFLTLSPIATACIAEKPVNNRFNFNSEQLAR
                    LRKARVNHWRDGDTLRSVSFANNHQKPIRIYAIDTPEKAVLSIQRKSEIELKEANKATE
                    FAKSLIPIGSEVWIWPLNSYSYDREVAAVFFKTNPLQLHFSFAVEMVANGHALPIAG
                    NDPDFVFSDLDPFNPLKIVGIELANGLNNAFNNRKNIFSYLENSFQSITMVYQQRGVD
                    QSWTRYLAPSNDFSSTKLGLGTIYELKLNNG"
gene     2009..3766
                    /gene="MG187"
CDS      2009..3766
                    /gene="MG187"
                    /note="similar to GB:U00089 SP:P75264 PID:1673667 percent
                    identity: 61.74; identified by sequence similarity;
                    putative"
                    /codon_start=1
                    /transl_table=4
                    /product="ABC transporter, ATP-binding protein"
                    /protein_id="AAC71406.1"
                    /db_xref="GI:3844784"
                    /translation="MANNKSAIELKNIVVDFGESVAIDNINLSVEKGHQLVSLLGPSGC
                    GKTTTLAVIAGLIKPTSGQVLFNGYDVTKKPPQERKIGLVFQNYALYPHMNVFENIVF
                    PLYSDNSWKQAVLEKNSVANHEINCLLLTSNGASVQEIDQLNKLFHDSIEKPKQIQYQ
                    INDLNVSVFKNLNELTANLKLIPSKHQFAITNLNKQTLKQINELEAEFKTKWKLQKQT
                    PIKSGVEHNAKLQAIKQHFSYEKQRLKKGHYFKTKVELKQTLVENLKLVKKAISEQTKL
                    IKQSSDYTKLQLKRLKIKVEPNQLKKQYKVPLNQLIKNYSLKTDKLTDTQLNEIEQIK
                    TRIVSIKQFINKTALEVANKLAITKLLTKPDKISGGQOQRVAIARAIVRRPKLLMD
                    EPLSNLDAKLRVQTRQWIRQQEBLQITVPVTHDQBEAMSIAGVTIVCMSTGKVQQIG
                    TPSELKLJKPANEFVARFLGTPEMNILECSVRKNLFHDLLVTESFKLNVEKLLVGD
                    RYEQLVVTTNKSSLQAKLINIENLGKHLVATISLFDTTLSWRLELNSHLKVGDSLNFI
                    IKANNLHFPDIDTKQRIEI"
gene     3769..4758
                    /gene="MG188"
CDS      3769..4758
                    /gene="MG188"
                    /note="similar to GB:U00089 SP:P75263 PID:1673666 percent
                    identity: 74.09; identified by sequence similarity;
                    putative"
                    /codon_start=1
                    /transl_table=4
                    /product="ABC transporter, permease protein"
                    /protein_id="AAC71407.1"
                    /db_xref="GI:3844785"
                    /translation="MFKWLLKSHNQPHSLQLGLLDQPLPFWKPFLLPALLTTLFT
                    IPFFLSLQKGFSANSDLYDLSSQSFSLRTFQDLFSESNFVLGLRNSFLYSLISILPFS
                    IIAIVIASAIVFVYKKLLRGFWQTVFPLFYVTSGVAISIAFVYIFDSASGILNTVFN
                    VNTKWLDSGSRDTFNALWAILFGVWKNLAFNVLIISTAMLSVNPCLYKVASLDSANP
                    VRQFFKITLPSIRPTLIFPLTTLLILGGQGATSDSDTBKISFKPGNQIDFN
                    IQSGNTNLAGAATLVLFVLGVCYGLVLRNGFYLIEWLQWKIKQLYVQKQLTLY"
gene     4682..5686
                    /gene="MG189"
CDS      4682..5686
                    /gene="MG189"
                    /note="similar to GB:U00089 SP:P75262 PID:1673665 percent
                    identity: 69.67; identified by sequence similarity;
                    putative"
                    /codon_start=1
                    /transl_table=4
                    /product="ABC transporter, permease protein"
                    /protein_id="AAC71408.1"

                    /db_xref="GI:3844786"
                    /translation="MAFIWLNGCSEKLNSFMFKNNLRFTSWINQHKFYQLDLSLKTRS
                    IKQLVLTLVFKTLVLGPGLIVIPPFYLMVVVSFASDERALDTRTPILMPDSWNFDNF
                    SRVLSDGKYLNAIVVNTLVTVLSVLTLLPFTICMGYSFSLRKWKYKKLVWFFLSVLI
                    LPESALLIGQYRIVIVANWNNPNSPLIVGLIMPFVSSVFSGFMYRTSFAIPSQLKB
                    SALIDGCNGFNYFLKIALPMVKSTSWTVGILTAFSAWNSYLWPLLLGNRVDLNINLW
                    VLQQGILDANSSDEQIRTLNLKMSAAILAILPMFIIYFLFHKRIMNAIKNRANTIKG"
gene     5726..6811
                    /gene="MG190"
CDS      5726..6811
                    /gene="MG190"
                    /note="similar to PID:1196657 GB:U00089 SP:P75144
                    PID:1673660 percent identity: 60.38; identified by
                    sequence similarity; putative"
                    /codon_start=1
                    /transl_table=4
                    /product="mgpA protein (mgpA)"
                    /protein_id="AAC71409.1"
                    /db_xref="GI:3844787"
                    /translation="MGKTSTISNHCVIRDYLLSVCHLWHCASDHYHYQWRGTFIFVM
                    KKGSTEAINAIKQFDKIVIFHHVRPDGDCLGAQQLFLHLIKANFKNKEVKCVGNNNN
                    LFSPINMTFNQIDESFLKEALAIVDDANYKNRIELRELLDKNLFKAVLRIDHHPNED
                    DLNTSFNFVBESYVACCEQIVEMATVAKWTIPPVAATLLYIGIYTDSNRFLYSNTSYR
                    TLYLAAILYKAKADIRIVHDHLNHTSLADLKFKKCVYNHFKTCQQVIYFICTKKIQKR
                    LRMTADQCARVNLLSNIADYKIWLFFIEQANNEIRIDLRSNGINVRDIAIKYGGGHN
                    NASGAIITNKKQISDVVSDCVKKIVYN"
gene     6818..11152
                    /gene="MG191"
CDS      6818..11152
                    /gene="MG191"
                    /note="similar to SP:P20796 GB:M31431 GB:U01694 GB:U01779
                    GB:U02159 percent identity: 98.89; identified by sequence
                    similarity; putative"
                    /codon_start=1
                    /transl_table=4
                    /product="adhesin P1 (mgpB)"
                    /protein_id="AAC71410.1"
                    /db_xref="GI:3844788"
                    /translation="MHQPKCRLAKKSWAFLTAALTLGVITGVGYFLFNQNKQRSSVS
                    NFAYQPKQLSVKHQQAVDETLTPWTWNNNFSSLKITGENPGSFGLVRSGNDNLNISS
                    VTKNSSDDNLKYLNAVEKYLDGQQNFAIRRYDINLAKMENPSTVQRGLNG
                    BPIFDPFKGPGLTGNAPTDWNEIKGKVPEVVQSPHSPNLYFVLLVPKVALEYHNLNN
                    QVVKESLLEVKATQSSFNPTQRLQKDSYKVDSSKQGEKLSETTASSMSSGMATSTRAKA
                    LKVEVERGSQSDSLLKNDFAKKPLKHKNSSGEBVKJEAEKEFTEAWKPLLTTDQIAREK
                    GMGATVVSFYDAPYSENHTAFGLVDHIDPKKMVENYPPSWKTPKWNHHGIWDYNARNL
                    LLQTTGFFNPRRHPEWFDREGQAKADNTSPGFKVGDTDHKDGFKNSSSPIALPPEAY
                    FANIGNMVAIGNSVFIPGGNGHATKMFTNPLSIGVPRIKYTDNPSKSSVTGWPYAVL
                    FGGLINPQTNGLKDLPLGTNRWFEYVPRMAVSGVKWVGNQLVLAGTLIMGDTATVPRL
                    KYDQLEKHLNLVAQOGGLLREDLQFTPYGWANRPDIPVGAWLQDEMGSKFGPHYFLN
                    NPDIQDNVANDTVBALISSYKNTDKLKHVYPYRYSGLYAWQLFNWSNKLTNTPLSANF
                    VNENSYAPNSLFAAILNEDLITGLSDKIFYGKENBFAENEADRFNQLLSLNPNNTNW
                    ARYLNVVQRFTTGPNLDSSTFDQFLDFLPWINGKFPFSNSPSPSTSASSSTPLPTFSN
                    INVGVKSMITQHLNKENTRWVFIPNFSPDITWTGAGYRVQSANQKNGIPPEQVKPSNNS
                    TPFDPNSDDNKVTPSGGSSKPTTVPALPNSISPTSDWINALTFTNKNNPQRNQLLRS
                    LLGTIPVLINKSGDSNDQFNKDSEQKWDKTETNEGNLPGFGEVNGLYNAALLHTYGFF
                    GTNTNSTDPKIGFKADSSSSSSTIVGSGLNWTSQDVGNLVVINDTSFGLQGWFIT
                    PTDFIRPRTGYLGTTLSLQDQTIIWADQPWTSPKGSYLDSDGTPLSNWDPTALKSLP
                    NSSTTYDTNPTLSPSPQLYQPNKVKAYQTNTYNKLIEPVDATSAATNMTSLLKLLTT
                    QNIKAKLGKTASSCGNNNGGGVSQTINTITTTGNISEGLKEETSIQAETLKKFPDSK
                    QNNKSBIGIGDSTFKMDGGLTGVVSTNCLLNLINGQGATSDSDTBKISFKPGNQIDFN
                    RLFTLPVTELFDPNTMFVYDQYVPLLVNLPSGFDQASIRLKVISYSVENQTLGVRLBP
                    KDPQTQQFIPVLNASSTGPQTVFQPFNQWADYVLPLIVTVPIVVIILSVTLGLTIGIP
                    MHRNKKALQAGFDLSNKKVDVLTKAVGSVFKEIINRTGISNAPKLKQATPTKPTKT
                    PPKPFVKQ"
gene     11154..14312
                    /gene="MG192"
CDS      11154..14312
                    /gene="MG192"
                    /note="similar to SP:P22747 GB:U02157 GB:M31431 GB:U02161
                    GB:X61522 percent identity: 99.43; identified by sequence
                    similarity; putative"
                    /codon_start=1
```

```
RESULT 13
MYCATTP
LOCUS       MYCATTP              9691 bp    DNA     linear   BCT 15-FEB-1996
DEFINITION  M.pneumoniae P1 operon encoding P1 surface protein, ORF-4, and
            ORF-6 proteins, complete cds.
ACCESSION   M21519 M20916
VERSION     M21519.1  GI:150138
KEYWORDS    P1 attachment protein; P1 surface protein.
SOURCE      Mycoplasma pneumoniae
  ORGANISM  Mycoplasma pneumoniae
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1054 to 5937)
  AUTHORS   Inamine,J.M., Denny,T.P., Loechel,S., Schaper,U., Huang,C.H.,
            Bott,K.F. and Hu,P.C.
  TITLE     Nucleotide sequence of the P1 attachment-protein gene of Mycoplasma
            pneumoniae
  JOURNAL   Gene 64 (2), 217-229 (1988)
  MEDLINE   88297153
  PUBMED    2841195
REFERENCE   2  (bases 1 to 9691)
  AUTHORS   Inamine,J.M., Loechel,S. and Hu,P.C.
  TITLE     Analysis of the nucleotide sequence of the P1 operon of Mycoplasma
            pneumoniae
  JOURNAL   Gene 73 (1), 175-183 (1988)
  MEDLINE   89211947
  PUBMED    2468577
COMMENT     Original source text: Mycoplasma pneumoniae (strain M129) (tissue
            library: ATCC 29342) DNA.
            Draft entry and computer-readable sequence for [2] kindly provided
            by S.Loechel, 16-NOV-1988.
FEATURES             Location/Qualifiers
     source          1..9691
                     /organism="Mycoplasma pneumoniae"
                     /mol_type="genomic DNA"
                     /strain="M129"
                     /db_xref="taxon:2104"
                     /tissue_lib="ATCC 29342"
     mRNA            54..9616
                     /product="P1 operon mRNA"
     CDS             316..1041
                     /note="ORF-4; putative"
                     /codon_start=1
                     /transl_table=4
                     /product="unknown protein"
                     /protein_id="AAA88324.1"
                     /db_xref="GI:1196657"
                     /translation="MQQALAVIVDANYKERIECRDLLDQNQFKAVLRIDHHPNEDDLN
                     TTHNFVDASYIAAAEQVVDLAVQAKWKLSPPAATALYLGIYTDSNRFLYSNTSWRTLY
                     LGSMLYRAQANIAKIHDELNHTSLKDIQFKQYVFKNFQTFQNVIYFVADKKFQKKLKV
                     TPLECARVNILANIEQFHIWLFFIBEGKNHYRVBFRSNGINVREVALKYGGGGHIQAS
                     GAVLKSKRDIIRVVQDCQKQIAV"
     CDS             1054..5937
                     /codon_start=1
                     /transl_table=4
                     /product="P1 protein"
                     /protein_id="AAA88325.1"
                     /db_xref="GI:150139"
                     /translation="MHQTKKTALSKSTWILILTATASLATGLTVVGHFTSTTTLKRQ
                     QFSYTRPDFVALRHTNAINPRLTPWTYRNTSFSSLPLTGENPGAWALVRDNSAKGITA
                     GSSGSQQTTYDPTRTEAALTASTTFALRYDLAGRALYDLDFSKLNPQTPTRDQTGQIT
                     FNPPGGFGLGSGAAPQQWNEVKNKVFVEVVAQDPSNPYRPAVLLVPRSVVYVEQLQRGLG
                     LPQQRTESQQNTSTGAMFGLKVVNABADTAKSNEKLQGABATGSSTTSGSGQSTQRG
                     GSSGDTKVALKIBVKKKSDSEDNGQLQLEKNDLANAPIKRSEBSGSVQLKADDFGT
                     ALSSSGGGNSNPGSPTPWRPWLATEQIHKDLPKWSASILILYDAPYARNRTAIDRVD
                     HLDPKAMTANYPPSWRTPKWNNHGLWDWKARDVLQTTGFFNPRRHPENFDGGQTVAD
                     NEKTGPDVDNSENTKQGPQKEADSDKSAPIALPFEAYPANIGNLTWFGQALLVPGGNG
                     HVTKSAHTAPLSIGVFRVRVNATGSAATVTGWPYALLFSGMVNKQTDGLKDLPNNNR
                     WFEYVPRMAVAGAKFVGRELVLAGITMGDTATVPRLLYDELBSNLNLVAQGGLLRE
                     DLQLFTPYGWANRPDLPIGAWSSSSSSSHNAPYYFHNNPDWQDRPIQNVVDAFIKPWE
                     DKNGKDDAKYIYPYRYSGMWAWQYYNWSNKLTDQPLSADFVNENAYQPNSLFAAILNP
                     ELLAALPDKVKYGKENEFAANEYERFNQKLTVAPTQGTNWSHFSPTLSRFSTGFNLVG
```

Query Match   3.9%;   Score 124.6;   DB 1;   Length 15787;
Best Local Similarity   63.5%;   Pred. No. 1.9e-11;
Matches 190;   Conservative 0;   Mismatches 109;   Indels 0;   Gaps 0;

```
Qy    2770 TTTGGGCCGCCTTACCTTCATGGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
           ||||| |||||||||  ||||| ||||||  |||||||||||  |||| ||  ||||||||||
Db   13950 TTTGCAGCACTCCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGGATCTTG 14009

Qy    2830 GCAATTATGATCATCTTAGGATTAGCTATCCGTATTCCTTTAAGAGCTCAAAGAAAATTA 2889
           |||||||||||||||||||||| |||| ||||| |||    ||| |||| ||| ||||||
Db   14010 TTTATCTTGTTAGTCTTAGTAGTCTTGGGATTGGGATCCCAATGTACACGGGTAAGAAACTC 14069

Qy    2890 CAAGACAAAGGGTTCAAAACAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT 2949
           ||||| || |||||||||| |||| |||||||||||| ||||||| ||||| |||||| |||
Db   14070 CAAGATGCATCGTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAACTGCTGCTGTCGGT 14129

Qy    2950 TCAGTTTACAAGAAGAATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTGCTTTA 3009
           |||||||  ||||||||| |||||| |||||||||||||||  ||||||||||||||| ||
Db   14130 AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGGTGAAAAAAGCACCTAGTGCATTG 14189

Qy    3010 GGTGCTCGTAAATCTGGTGATAAGAAACCTGCTGCTCGTCGCTAAACCACCTGTTCAACC 3068
           |||||||||||||||| ||||||||||||||||||||| |||  |||||||||||||||
Db   14190 AAAGCTGCTAATCCTAGTGTGTAAAAAAACCTGCTGCTGCTTTTTTAAAACCACCTGTTCAACC 14248
```

RESULT 12
AR300198_2
WPCOMMENT
Sequence split into 6 fragments   LOCUS AR300198 Accession AR300198

Fragment Name	Begin	End
AR300198_0	1	110000
AR300198_1	100001	210000
AR300198_2	200001	310000
AR300198_3	300001	410000
AR300198_4	400001	510000
AR300198_5	500001	580073

Continuation (3 of 6) of AR300198 from base 200001 (AR300198 Sequence 1 from patent US 6

Query Match   3.9%;   Score 124.6;   DB 6;   Length 110000;
Best Local Similarity   63.5%;   Pred. No. 1.2e-11;
Matches 190;   Conservative 0;   Mismatches 109;   Indels 0;   Gaps 0;

```
Qy    2770 TTTGGGCCGCCTTACCTTCATGGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
           ||||| |||||||||  ||||| ||||||  |||||||||||  |||| ||  ||||||||||
Db   28703 TTTGCAGCACTCCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGGATCTTG 28762

Qy    2830 GCAATTATGATCATCTTAGGATTAGCTATCCGTATTCCTTTAAGAGCTCAAAGAAAATTA 2889
           |||||||||||||||||||||| |||| ||||| |||    ||| |||| ||| ||||||
Db   28763 TTTATCTTGTTAGTCTTAGTAGTCTTGGGATTGGGATCCCAATGTACACGGGTAAGAAACTC 28822

Qy    2890 CAAGACAAAGGGTTCAAAACAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT 2949
           ||||| || |||||||||| |||| |||||||||||| ||||||| ||||| |||||| |||
Db   28823 CAAGATGCATCGTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAACTGCTGCTGTCGGT 28882

Qy    2950 TCAGTTTACAAGAAGAATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTGCTTTA 3009
           |||||||  ||||||||| |||||| |||||||||||||||  ||||||||||||||| ||
Db   28883 AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGGTGAAAAAAGCACCTAGTGCATTG 28942

Qy    3010 GGTGCTCGTAAATCTGGTGATAAGAAACCTGCTGCTCGTCGCTAAACCACCTGTTCAACC 3068
           |||||||||||||||| ||||||||||||||||||||| |||  |||||||||||||||
Db   28943 AAAGCTGCTAATCCTAGTGTGTAAAAAAACCTGCTGCTGCTTTTTTAAAACCACCTGTTCAACC 29001
```

/transl_table=4
/product="mgpC protein (mgpC)"
/protein_id="AAC71411.1"
/db_xref="GI:3847489"
/translation="MKTNRKQIYKKAYWLLLPFLPLALANTFLVKEDSKNVTAYTPFA
TPITDSKSDLVSLAQLDSSYQIADQTIHVTNLFVLFKSRDVKVKYBSSGSNNISFDST
SQGEKPSYVVEFTNSTNIGIKWTMVKKYQLDVPNVSSDMNQVLRNLILEQPLTKYTLN
SSLAEEKGKTQREVHLGSQANQWTSQRNQHDLNNPSPNASTGFKLTTGNAVRKLSE
SWPIYEPIDGTKQGKGKDSGWSSTEENEAKNDAPSVSGGGSSSGTFNKYLNTKQALE
SIGILFDDQTPRNVTLQLYIASTSKLAVTNNHIVVMGNSFLPSTWYWVVERSAQENAS
NKPTWFANTNLDWGEDKQKQFVENQLGYKETTSTNSHNFHSKSFTQPAYLISGIDSVN"

```
SVLDQVLDYVPWIGNGYRYGNHRGVDDITAPQTSAGSSSGISTNTSGGRSFLPTPSN
IGVGLKANVQATLGGSQTMITGGSPRRTLDQANLQLWTGAGWRNDKASSGQSDENHTK
FTSATGMDQQQSQGSGTSAGNPDSLKQDNISKSGDSLTTQDGNAIDQQEATNYTNLPPNL
TPTADHPNALSPTNKNNAQRAQLFLRGLLGSIPVLVNRSGSDSNKFQATDQKWGYTDL
HSDQTKLNLPAYGEVNGLLNPALVETYFGNTRAGSGSNTTSSPGIGFKIPEQNNDSK
ATLITPGLAWTPQDVGNLVVSGTTVSFQLGGWLVTFTDFVKPRAGYLGLQLTGLDASD
ATQRALIWAPRPWAAFRGSWNRLGRVESVWDLKGVWADQAQSDSQGSTTATRNALP
EHPNALAFQVSVVEASAYKPNTSSGGTQSTNSSPYLHLVKPKKVTQSDKLDDDLKNLL
DPNQVRTKLRQSFGTDHSTQPQPGSLGKTTTPVFGTSSGNLSSVLSGGAGGGSSGSGQ
SGVDLSPVEKVSGWLVGQLPSTSDGNTSSTNLAPTNTGNDVVGVRLSESNAAKMN
DDVDGIVRTPLAELLDGEGQTADTGPQSVKFKSPDQIDFNRLFTHPVTLFDPVTMLV
YDQYIPLFIDIPASVNPKMVRLKVLSPDTNEQSLGLRLEFFKPDDTQPNNNVQVNPN
NGDFLPLLTASSQGPQTLFSSPNQWPDYVLPLAITVPIVVIVLSVTLGLAIGIPHHKN
KQALKAGFALSNQKVDVLTKAVGSVFKEIINRTGISQAPKRLKQTSAAKPGAPRPPVP
PKPGAPKPPVQPPKKPA"
```

CDS   5943..9599
                /note="ORF-6; putative"
                /codon_start=1
                /transl_table=4
                /product="P1 protein"
                /protein_id="AAA88326.1"
                /db_xref="GI:1196658"
                /translation="MKSKLKLKRYLLFLPLLPLGTLSLANTYLLQDHNTLTPYTPFTT
PLNGGLDVVRAAHLHPEYELVDWKRVGDTKLVALVRSALVRVKFQDTTSSDQSNTNQN
ALSPDTQESQKALNGSGSGSDTSGSNSQDFASYVLIFKAAPRATWVFERKIKLALPY
VKQESQSGDGQSNGKGSLYKTLQDLLVEQPVTPYTPNAGLARVNGVAQDTVHEGSGQ
ESSWNSQRSQKGLKNNPGFKAVTGFKLDKGRAYRKLNESWPVYEPLDSTKEGKDES
SWKNSEKTTAENDAPLVGMVSGGAAGSASSLQGNGSNSSGLKSLLRSAPVSVPPSSTS
NQTLSLSNPAPVGPQAVVSQPAGGATAAVSVNRTASDTAIFSKYLNTAQALHQMGVIV
PGLEKWGGNNGTGVVASRQDATSTNLPHAAGASQTGLGTGSPREPALTATSQRAVTVV
AGPLRAGNSSETDALPNVITQLYHTSTAQLAYLNGQIVVMGSDRVPSLWTWVVGEDQE
SGKATWWAKTELNWGTDKQKQFVENQLGFKODSNDSKNSNLKAQGLTQPAYLIAGLD
VVADHLVFAAFKAGAVGYDMTTDSSASTYNQALAWSTTAGLDSDGGYKALGVENTAGLN
GPINGLPTLLDTFAYVTPVSGMKGGSQNNEEVQTTYPVKSDQKATAKIASLINASPLN
SYGDDGVTVPDALGLNFNFKLNEERLPSRTDQLLVVGIVNESELKSARENAQSTSDDN
SNTKVKWTNTASHYLPVPYYISANFPEAGNRRAEQRNGVKISTLESQATDGFANSLL
NFGTGLKAGVDPAPVARGHKPNYSAVLLVRGVVRLNFNPDTDKLLDSTDKNSEPISF
SYTPFGSABSAVDLTTLKDVTYIAESGLWFYTFDNGEKPTYDGKQQVKNRKGYAVIT
VSRTGIEFNEDANTTTLSQAPAALAVQNGIASSQDDLTGILPLSDEFSAVITKDQTWT
GKVDIYKNTNGLFEKDDQLSENVKRRDNGLVPIYNEGIVDIWGRVDFAANSVLQARNL
TDKTVDEVINNPDILQSPFKFTPAFDNQRAMLVGEKTSDTTLTVKPKIEYLDGNFYGE
DSKIAGIPLNIDFPSRIFAGFAALPSWVIPVSGSSVGILLLILGLGIGIPMYKVR
KLQDSSFVDVFKKVDTLLTTAVGSVYKKIITQTSVIKKAPSALKAANNAAPKAPVKPAA
PTARPPVQPPKKA"

terminator   9617..9651
                /note="putative"
ORIGIN   526 bp upstream of HindIII site.

Query Match   3.2%;   Score 102.2; DB 1; Length 9691;
Best Local Similarity 55.5%;   Pred. No. 1.3e-07;
Matches 197; Conservative   0; Mismatches 158; Indels   0; Gaps   0;

Qy   2737 AACAATAAATTAGTAACATTCTTCTTATAACAACTTGGCGCCTTACCTTCATGGGTAGTG 2796
            ||| |||| |||| |||||| ||| || ||||||| || ||||||||| ||||
Db   9234 AACATTGATTTCCCTTCCCGGATTTTGCTGGCGTTTACCGTCCTCGGGTCATT 9293

Qy   2797 CCTACAGCAATTGGTAGTACATTAGGTATTCTTGCAATTATGATCATCTTAGGATTAGCT 2856
            ||| |||| |||||| ||| |||||| ||| ||||||| ||||| ||||| |||||
Db   9294 CCGGTATCAGTCGGTTCATCGGTGGCATTCTCTTAAATCCTGCTCATCTTAGGCCTTGGT 9353

Qy   2857 ATCGGTATTCCTTTAAGAGCTCAAAGAGAAATTACAAGACAAAGGGTTCAAAACAACATTC 2916
            ||| |||| ||| |||||
Db   9354 ATTGGAATTCCAATGTATATAGGTCCGCAAGCTTCAAGACTTCCAGCTTTGTTGATGTGTT 9413

Qy   2917 AAAAAAGTTGATAACCTGACTGCGTGCTGTTGGTTCAGTTACAAGAAGATTATTACCCCAA 2976
            |||||| ||| ||| | |||||||||||| || ||||||||| |||||||||||||||
Db   9414 AAAAAGGTGGATAACGTTGACAACCGCTGTGGTAGCGTGTACAAGAAGATTATCACCCAA 9473

Qy   2977 ACTGCTAACGTTAAGAAGAAAAAACCTGCTCGTTAGGTAGCTGGTAAATCTGGTGATAAGAAA 3036
            ||| ||||||| | ||||||
Db   9474 ACGAGTGTGATCAAAAAAGTCCCTAGTCCGTTGAAAGCTGCTAATAACGCTGCTCCTAAA 9533

Qy   3037 CCTGCTGCTAGCTGCTAAACCTGCTCCCAGCTAAACCATCTGCACCAAAAGCTA 3091
```

Db   9534 GCACCAGTTAAACCAGCTGCTCCAACAGCTCCAAGACCACCAGTCCAACCACCTA 9588
                      |||| | ||| ||| | ||| ||||| ||| ||| | |||| |||||

Db   9534

RESULT 14
AE000002

LOCUS   AE000002   16876 bp   DNA   linear   BCT 24-NOV-2000
DEFINITION Mycoplasma pneumoniae M129 section 15 of 63 of the complete genome.
ACCESSION   AE000002 U00089
VERSION   AE000002.2 GI:11379441
KEYWORDS   .
SOURCE   Mycoplasma pneumoniae
  ORGANISM Mycoplasma pneumoniae
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 16876)
  AUTHORS   Himmelreich,R., Hilbert,H., Plagens,H., Pirkl,E., Li,B.C. and
            Herrmann,R.
  TITLE   Complete sequence analysis of the genome of the bacterium
            Mycoplasma pneumoniae
  JOURNAL   Nucleic Acids Res. 24 (22), 4420-4449 (1996)
   MEDLINE   97105885
    PUBMED   8948633
REFERENCE   2 (bases 1 to 16876)
  AUTHORS   Dandekar,T., Huynen,M., Regula,J.T., Ueberle,B., Zimmermann,C.U.,
            Andrade,M.A., Doerks,T., Sanchez-Pulido,L., Snel,B., Suyama,M.,
            Yuan,Y.P., Herrmann,R. and Bork,P.
  TITLE   Re-annotating the Mycoplasma pneumoniae genome sequence: adding
            value, function and reading frames
  JOURNAL   Nucleic Acids Res. 28 (17), 3278-3288 (2000)
   MEDLINE   20411492
    PUBMED   10954595
REFERENCE   3 (bases 1 to 16876)
  AUTHORS   Himmelreich,R., Hilbert,H. and Li,B.-C.
  TITLE   Direct Submission
  JOURNAL   Submitted (15-NOV-1996) Zentrun fuer Molekulare Biologie
            Heidelberg, University Heidelberg, 69120 Heidelberg, Germany
REFERENCE   4 (bases 1 to 16876)
  AUTHORS   Suyama,M., Dandekar,T. and Herrmann,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (15-JUN-2000) Zentrun fuer Molekulare Biologie
            Heidelberg, University Heidelberg, 69120 Heidelberg, Germany
COMMENT   On Nov 27, 2000 this sequence version replaced gi:1673651.
            This updated annotation replaces the old annotation from reference
            1. The old gene identifiers (MP numbers) according to the original
            publication by
            Himmelreich et al. (1996) are given as well as new gene numbering
            (MPN numbers) from the origin of replication. Annotation comments
            and further update data are at
            http://www.bork.embl-heidelberg.de/Annot/MP/.
FEATURES   Location/Qualifiers
     source   1..16876
                     /organism="Mycoplasma pneumoniae"
                     /mol_type="genomic DNA"
                     /strain="M129"
                     /db_xref="taxon:2104"
                     /note="ATCC 29342"
     gene   138..1112
                     /gene="orf4"
                     /note="synonym: B07_orf324"
     CDS   138..1112
                     /gene="orf4"
                     /note="MPN140(new), 015 (Himmelreich et al., 1996)"
                     /codon_start=1
                     /transl_table=4
                     /product="DHH family phosphoesterases"
                     /protein_id="AAB95662.1"
                     /db_xref="GI:1673660"
                     /translation="MNSQVRKGSIABAVSAIQAHDKIVIFHHIRPDGDCLGAQHGLA
RLIQTNFPHKQVFCVGDPKENFPWLEMVFKEQITPELMQQALAVIVDANYKERIEC
RDLLDQNQFKAVLRIDHHPNEDDLNTTHNFVDASYIAAARQVDLAVQAKWKLSPPAA
TALYLGIYTDSNRPLYSNTSWRTLYLGSMLYRAQANIAKIHDBLNFTSLKDIQFKQYV
FKNFQTFQNVIYFVADKKFQKKLKVTPLECARVNILANIEQFHIHLFPIEGKNHYRV
EFRSNGINVREVALKYGGGGHIQASGAVLKSKRDIIRVVQDCKQIAV"
```

```
gene            1125..6008
                /gene="P1"
                /note="synonym: E07_orf1627"
CDS             1125..6008
                /gene="P1"
                /note="MPN141(new), 014(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="ADP1 MYCPN adhesin P1"
                /protein_id="AAB95661.1"
                /db_xref="GI:1673659"
                /translation="MHEQTKKTALSKSTWILILTATASLATGLTVGHFTSTTTLKRQ
                QFSYTRPDEVALRHTNAINPRLITPWTYRNTSFSSLPLTGENPGAWALVRDNSAKGITA
                GSGSQQTTYDPTRTEAALTASTTPALRRYDLAGRALYDLDPSKLNPQTPTRDQTQGIT
                FNPFGGFGLSGAAPQQWNEVKNKVPVEVAQDPSNPYRFAVLLVPRSVVYTRQLQRGLG
                LPQQRTBSGQNTSTTGAMFGLKVKNAEADTAKSNEKLQGAEATGSSTTSGSGQSTQRG
                GSSGDTVKALKIEVKKKSDSEDNGQLQEKNDLANAPIKRSEBSGQSVQLKADDPFGT
                ALSSSGSGGNSNPGSPTPWRPWLATEQIHKDLPKWSASILLIYDAPVARNRTAIDRVD
                HLDPKAMTANYPPSWRTPKWNHHGLWDWKARDVLLQTGFFNPRRHPEWFDGGQTVAD
                NEKTGFDVDNSENTKQGFQKEADSDKGSAPIALPPEAYPANICNLTWPGQALLVFGGNG
                HVTKSAHTAPLSIGVFRVRYNATGTSATVTGWPYALLFSGMVNKQTDGLKDLPFNNNR
                WFEIVPRMAVAGAKFVGRELVLAGTITMGDTATVPRLLYDELESNLNIVAQCGLLRE
                DLQLFFTPYGWANRPDLFIGAWSSSSSSHNAPYYFHNNPDWQDRPIQNVVDAFIKPWE
                DKNGKDDAKYIYPYRYSGMWAWQVYNWSNKLTDQPLSADFVNENAYQPNSLFAAILNP
                ELLAALPDKVKYGKENEPAANEYBRFNQKLITVAPTQGTNWSHFSPTLSRFSTGFNLVG
                IGVGLKANVQATLGGSQTMITGGSPRRTLDQANLQLWTGAGWRNDKASSGQSDENHTK
                FTSATGMDQQGQSGTSAGNPDSLKQDNISKSGDSLTTQDGNAIDQQEATNVTNLPPNL
                TPTADWPNALSFTNKNNAQRAQLFLRGLLGSIPVLVNRSGSDSNKFQATDQKWSYTDL
                HSDQTKLNLPAYGEVNGLLNPALVETYPNTRAGGSGSNTTSPGIGFKIPBQNNDSK
                ATLITPGLAWTPQDVGNLVVSGTTVSPQLGGWLVTFTDFVKPRAGVLGLQLTGLDASD
                ATQRALITHAPRPWAAFRGSWVNRLGRVESVWDLKGVWADQAQSDSQGSTTATRNALP
                EHPNALAFQVSVVEASAYKPNTSSGQTQSTNSSPYLHLVKPKKVTQGDKLDDDLKNLL
                DPNQVRTKLROSFGTDHSTQPQPQSLKTTTPVFGTSSGNLSSVLSGGGAGGSGSSGQ
                SGVDLSPVEKVSGWLVGQLPSISDGNISSTNNLAPNTNTGNDVVGVGRLSESNAAKVN
                DDVDGIVRTPLAELLDGEGQTAIDTGPQSVKFKSPDQIDFNRLIFTHTPGPLDPFNNRLFTHP
                YDQYIPLFFIDIPASVNPKMVRLKVLSPDTNEQSLGRLEPFKPDQTQPNNNVQVNPN
                NGDPLPLLTASSGPQOTLFSPFNQWPDVTVLPLAITVPIVIVLSVTLGLATGIPMHKN
                KQALKAGFALSNQKVDVLTKAVGSVFKIINRTGISQAPKRLKQTSAAKPGAPRPPVP
                PKPGAPKPVQPKKPA"
repeat_region   1125..2671
                /rpt_family="REPMP4"
repeat_region   3171..4971
                /rpt_family="REPMP2/3"
gene            6014..9670
                /note="synonym: orf6"
                /gene="E07_orf1218"
gene            6014..9670
                /gene="orf6"
gene            6014..9670
                /gene="orf6"
                /note="MPN142(new), 013(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="involved in cytadherence"
                /protein_id="AAB95660.1"
                /db_xref="GI:1673658"
                /translation="MKSKLKLKRYLLFLPLLPLGTLSLANTYLLQDRNTLTPYTPFT
                PLNGGLDVVRAAHLHPSYELVDWKRVGDTKLVALVRSALVRVKFQDTTSSDQSNTNQN
                ALSFDTQESQKALNGSGQSGSSDTSGSNSQDFASYVLIPKAAPRATWVPERKIKLALPY
                VKQESQGSGDDQSGNGKGSLYKTLQDLLVEQVTPYTPNAGLARVNGVAQDTVHFGSGQ
                ESSWNSQRSRQKGLKNNPGPKAVTGFKLDKGRAYRKLNESWPVVEPLDSTKEGKGKDES
                SWKNSEKTTAENDAPLVGMVSGAAGSASSLQGNGNSSGLKLLRSAPVSVPPSSTS
                NQTLSLSNPAPVGPQAVVSQPAGGATAAVSVNRTASDTATFSKYLNTAQALHQMGVIV
                PGLEKWGGNNGTGVVASRQDATSTNLPHAAGASQTGLGTGSPREPALTATSQRAVTVV
                AGPLRAGNSSETDALPNVITQLYHTSTAQLAYLNGQIVVMGSDRVPSLWYWVVGEDQE
                SGKATWAKTELNWGTDKQKQFVENQLGKDDSNSDKSNSNLKAQGLTQPAYLIAGLD
                VVADHLVPAAFKAGAVGYDMTTDSSASTYNQALAWSTTAGLDSDGYKALVENTAGLN
                GPINGLFTLLDTFAYVTPVSGMKGGSQNNEEVQTYPVKSDQKATAKIASLINASPLN
                SYGDDGVTVFDALGLNFNFKLNESELPSRTDLLVYGIVNESELKSARENAQSTSDDN
                SNTKVKWTNTASHYLPVPYYSANFPEAGNRRAEQRNGVKLISTLESQATDGFANSLL
                NFGTGLKAGVDPAPVARGHKPNYSAVLLVRGGVVRLNFNPDTDKLLDSTDKNSEPISF
                SYTPFGSABSAVDLITLKDVTYIAESGLWFYTFDNGEKPTYDGKQQYKNRKGYAVIT
```

```
repeat_region   6014..8714
                /rpt_family="REBPMP5"
gene            9787..10314
                /gene="E07_orf175"
CDS             9787..10314
                /gene="E07_orf175"
                /note="MPN143(new), 012(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="hypothetical protein"
                /protein_id="AAB95659.1"
                /db_xref="GI:1673657"
                /translation="MSHKDFNGLQAPQLLSSSSPVAKKQSHKLRHALKHARYLNHSS
                KRTLKHALELHEDNQVLLEBEGSPNFQDWLSKQPGVNKTSLKYNKSLGSWISKESKPK
                KRPPYFTYKGSKTTPEBAKALQQMKQSQKRFFHENMHSPLNEVAHNPMIQRFKQKQA
                KRAANTRQRTYKYRQ"
repeat_region   10457..11977
                /rpt_family="REBPMP2/3"
gene            10888..12129
                /gene="E07_orf413"
CDS             10888..12129
                /gene="E07_orf413"
                /note="MPN144(new), 011(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="hypothetical protein, see: MPN141"
                /protein_id="AAB95658.1"
                /db_xref="GI:1673656"
                /translation="MGQQGQSGTSAGNPDSLKQDKISKSGDSLTTQDGNATGQQEATN
                YTNLPPNLTPTADWPNALSFTNKNNAHRAQLFLRGLLGSIPVLVNRSGSDSNKFQATD
                QKWSYTDLQSDQTKLNLPAYGEVNGLLNPALVETYPGNTRAGGSGSNTTSPGIGFKI
                PEQNNDSKATLITPGLAWTPQDVGNLVVSGTSLSFQLGGWLVSFTDFIKPRAGVTGLQ
                LSGLDASDSDQRELIWAKRPWAAFRGSWVNRLGRVESVWDLKGVWADQAQLAAQAATS
                RASGSALAPHNALAFQVSVVEASAYSSSTSSGGSSSNTSPYLHLIKPKKVESTTQ
                LDQGLKNLLDPNQVRTKLRQSFGTDHSTQPQSLKTTTPVFGTSSGNIGSVLSGGGAGG
                GSSGSGSGVDLSPVERVSGH"
gene            12481..13020
                /gene="E07_orf179"
CDS             12481..13020
                /gene="E07_orf179"
                /note="MPN145(new), 010(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="hypothetical protein, see: MPN013"
                /protein_id="AAB95657.1"
                /db_xref="GI:1673655"
                /translation="MSYSPSLKBIIAILQKYTSKNYQSNCKTRPDGKLELSLNGVFBE
                IVKTPGKTRYVTHKQLDQKLKDFKQDLMVELHDTFATKTDLKESEARINQKLEALIQI
                VMVQGEQIKVEGEQINKLTQAVEKQGEKIBAQGQQIQKVNETLNFVVESLGSIHKRLD
                SMEGRLDSMENRLDKLBSK"
gene            13338..14135
                /gene="E07_orf265"
CDS             13338..14135
                /gene="E07_orf265"
```

```
Query Match      3.2%;   Score 102.2;  DB 1;   Length 16876;
Best Local Similarity  55.5%;   Pred. No. 1.1e-07;
Matches  197;   Conservative   0;   Mismatches  158;   Indels   0;   Gaps    0;

Qy    2737  AACATAATTAGTAACATTCTCTTATATACAACTTTGGCGCCTTACCTTCATGGGTAGTG   2796
            |||  ||    |  ||||  |      ||| |     |||| |||| || |||||||||||
Db    9305  AACATTGATTTCCCTTCCCGGATTTTGCTGGCTTTGCTGCTTACCGTCCTGGGTCATT   9364

Qy    2797  CCTCACGCAATTGGTAGTACATTAGGTATTCTTGCAATTATGATCATCTTAGGATTAGCT   2856
            |||    | |  |  |  ||  ||  | | ||   |||  |||  ||   |||| ||||||
Db    9365  CCGGTATCAGTCAGTCGGTTCATCGGTCGGGCATTCTTCTTAATCCTCGTCATCTTAGGCCTTGGT   9424
```

```
Qy   2857  ATCGGTATTCCTTTAAGAGCTCAAAGAAAATTACAAGACAAAGGGTTCAAAAACAACATTC  2916
            ||||| |||||||||||||||||| ||  ||| || |||| |||| |   |||
Db   9425  ATTGGAATTCCAATGTATAAGGTCCGCAAGCTTCAAGACTCCAGCTTTTGTTGTTGATGTGTTT  9484

Qy   2917  AAAAAAGTGATACCTTGACTGCTGTCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAA  2976
            |||| || |||| |||||  || ||||| |||||||| | |||||||||||||  |||||
Db   9485  AAAAAGGTGGATACGTTGACAACCGTGGGTAGCGTGTACAAGAAGATTATCACCCAA  9544

Qy   2977  ACTGCTAACGTTAAGAAAAAACCTGCTGCTTTAGGTGCGGTAAATCTGGTGATAAGAAA  3036
            ||    || ||| |||||||| |  || ||||| ||| |   || || |||| ||||||
Db   9545  ACGAGTGTGATCAAAAAAAGCTCCTAGTGCGTTGAAAGCTGCTATTAACCTGCTCCTAAA  9604

Qy   3037  CCTGCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTA  3091
            |  ||  ||||| |||| ||||||| |  |  |  ||| || || ||| ||
Db   9605  GCACCAGTTAAACCAGCTGCTCCAACAGCTCCAAGACCACCAACCACCACCTA  9659
```

RESULT 15
MGU34970

```
LOCUS       MGU34970                1618 bp    DNA     linear   BCT 22-FEB-1996
DEFINITION  Mycoplasma genitalium repetitive sequence element mgp-r5.
ACCESSION   U34970
VERSION     U34970.1  GI:1177840
KEYWORDS    .
SOURCE      Mycoplasma genitalium
  ORGANISM  Mycoplasma genitalium
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 1618)
  AUTHORS   Peterson,S.N., Hu,P.C., Bott,K.F. and Hutchison,C.A. III.
  TITLE     A survey of the Mycoplasma genitalium genome by using random
            sequencing
  JOURNAL   J. Bacteriol. 175 (24), 7918-7930 (1993)
  MEDLINE   94075230
   PUBMED   8253680
REFERENCE   2  (bases 1 to 1618)
  AUTHORS   Peterson,S.N., Bailey,C.C., Jensen,J.S., Borre,M.B., King,E.S.,
            Bott,K.F. and Hutchison,C.A. III.
  TITLE     Characterization of repetitive DNA in the Mycoplasma genitalium
            genome: possible role in the generation of antigenic variation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 92 (25), 11829-11833 (1995)
  MEDLINE   96102208
   PUBMED   8524858
REFERENCE   3  (bases 1 to 1618)
  AUTHORS   Peterson,S.N., Bailey,C.C., King,E.S., Bott,K.F. and Hutchison,C.A.
            III.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-1995) Clyde Hutchison, Microbiology and
            Immunology, The University of North Carolina, CB# 7290 F.L.O.B.,
            Chapel Hill, NC 27599-7290, USA
FEATURES             Location/Qualifiers
     source          1..1618
                     /organism="Mycoplasma genitalium"
                     /mol_type="genomic DNA"
                     /strain="G37"
                     /db_xref="taxon:2097"
                     /clone="eshl0"
     CDS             <1..270
                     /note="submitter believes protein coding sequences are not
                     expressed unless the sequence recombines into the
                     functional adhesin operon (MgPa); ORF3 protein"
                     /codon_start=1
                     /transl_table=4
                     /protein_id="AAA88899.1"
                     /db_xref="GI:1177841"
                     /translation="IPMYRVRKLQDASFVNVFKKVDDFLTTAVGSVYCKIITQTGVVKK
                     APSALKAANPSVKKPAAFLKPPVQPPSKPGEGQKAVEVKSEETKS"
     CDS             <655..>1068
                     /note="submitter believes protein coding sequences are not
                     expressed unless the sequence recombines into the
                     functional adhesin operon (MgPa); mgp-r5; similar to B
                     region of MgPa adhesin"
                     /codon_start=1
                     /transl_table=4
                     /protein_id="AAA88900.1"
                     /db_xref="GI:1177842"
                     /translation="LLVPKVVLEYHKLSKDVVKESLEVEATDSFDTQRLQKDSPMKD
                     SSKQGEKLSTTASSMSGMATSPRKALKIEVEKGSGSDTLTKSDFAKKPLKHKENSG
                     TEVKLDAQKDFAGEKAWKPLLTTDQIAREKGMGAT"
     CDS             1068..>1618
                     /note="submitter believes protein coding sequences are not
                     expressed unless the sequence recombines into the
                     functional adhesin operon (MgPa); mgp-r5; similar to BF
                     region of MgPa adhesin"
                     /codon_start=1
                     /transl_table=4
                     /protein_id="AAA88901.1"
                     /db_xref="GI:1177843"
                     /translation="LDFLPWIGNNKPFSNSHTASLSVSSNTPLFTFSNINVGVKSMIT
                     QHLNKENRWVFTPNSSPDIWTGAGYRKQGNNNGIPFDNVKPSNSSTPFNPNSDDNKVT
                     SGGSSKPTTYTHLPNSISPTSDWSNALTFTNKNNPQRNQLLRSLLGTIPVLINKSGT
                     GDQFNKDSEQKWNETDKLGGNLPG"
ORIGIN

Query Match             2.9%;  Score 93;  DB 1;  Length 1618;
Best Local Similarity  68.3%;  Pred. No. 6.7e-06;
Matches  129;  Conservative   0;  Mismatches  60;  Indels   0;  Gaps   0;

Qy   2880  AAGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTGATACCTTGACTGC  2939
            ||||||| || ||||   | ||||   |||||||   ||||||||||| ||  ||
Db     18  AAGAAAACTCCAAGATGCATCGTTGTTAATGTTCTTTAAAAAGGTTGATACACTCACAAC  77

Qy   2940  TGCTGTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACC  2999
            ||||||  |  ||| | || |||||||||| |||||||||||||| || ||||||| ||
Db     78  TGCTGTCGGTAGTGTGTACAAAAAGATTATTATCCAAACTGCTAACGTGTGTGAAAAAGCACC  137

Qy   3000  TGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTCCTGCTAAACCTGC  3059
            | ||||| | || |   ||||   || ||| ||||||||| |||| |  |    |||
Db    138  TAGTGCATTGAAAGCTGCTAARCCTAGTGTTAAAAAACCTGTTTAAAAAAACCTGCTTTTTTAAAACCACC  197

Qy   3060  TGCTCCAGC  3068
            ||| ||
Db    198  TGTTCAACC  206

Search completed: May  5, 2004, 17:06:54
Job time : 11960.1 secs
```

```
                  Copyright   GenCore version 5.1.6
                            (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   May  5, 2004, 09:54:21 ; Search time 1116.68 Seconds
                                   (without alignments)
                                   12131.925 Million cell updates/sec

Title:       US-09-901-572A-2
Perfect score:  3189
Sequence:    1 atgaatatttctaaaaaact.........caaccgctccaaagaataa 3189

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:    3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:       6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
             1: geneseqn1980s:*
             2: geneseqn1990s:*
             3: geneseqn2000s:*
             4: geneseqn2001as:*
             5: geneseqn2001bs:*
             6: geneseqn2002s:*
             7: geneseqn2003as:*
             8: geneseqn2003bs:*
             9: geneseqn2003cs:*
            10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result NO.	Score	% Query Match	Length	DB	ID	Description
	1	3189	100.0	3189	7	ACF03365	Acf03365 Mycoplasm
	2	3144.2	98.6	3189	2	AAT75087	Aat75087 DNA encod
	3	3143.4	98.6	8354	7	ABX95108	Abx95108 DNA encod
	4	3143.4	98.6	10651	6	AAF88025	Aaf88025 Vector pc
	5	3057.8	95.9	3129	7	ACF03435	Acf03435 Prokaryot
	6	124.6	3.9	3159	2	ACA39446	Aca39446 Prokaryot
	7	124.6	3.9	110000	2	AAT58840_2	Continuation (3 of
	8	78.4	2.5	4182	2	AAQ81778	Aaq81778 DNA encod
	9	64.6	2.0	3435	2	AAQ79746	Aaq79746 Mycoplasm
c	10	62.6	2.0	4985	6	ABQ75107	Abq75107 Anopheles
c	11	62.6	2.0	4985	9	ACF79720	Acf79720 Mosquito
c	12	60.8	1.9	8771	6	ABL33825	Abl33825 Human imm
c	13	60.2	1.9	3489	3	AAA30290	Aaa30290 Kaposi's
	14	60.2	1.9	3489	4	AAF82901	Aaf82901 Nucleotid
	15	60.2	1.9	3489	6	ABA93487	Aba93487 Kaposi's
	16	60.2	1.9	32207	2	AAV73805	Aav73805 KSHV LUR
	17	60.2	1.9	137507	2	AAV19941	Aav19941 KSHV long
c	18	59	1.9	535	6	ABQ42292	Abq42292 Oligonucl
	19	59	1.9	535	6	ABQ42293	Abq42293 Oligonucl
	20	58.4	1.8	110000	6	ABA92787_5	Continuation (6 of
c	21	58	1.8	50000	6	ABL55644	Abl55644 AmEPV gen
c	22	55.8	1.7	969	6	ABQ39490	Abq39490 Oligonucl
c	23	55.8	1.7	969	6	ABQ39491	Abq39491 Oligonucl
c	24	55	1.7	2000	7	ADA71938	Ada71938 Rice gene
	25	54.8	1.7	3837	3	AAA70211	Aaa70211 Plasmodiu
	26	54.6	1.7	14041	4	AAH48024	Aah48024 Internal
	27	53.8	1.7	3582	3	AAA70241	Aaa70241 Plasmodiu
	28	53.4	1.7	1159	3	AAA59240	Aaa59240 An EcoRI
	29	53.2	1.7	987	3	AAT90555	Aat90555 Clostridi
	30	53.2	1.7	1037	3	AAA59242	Aaa59242 Exons E,
	31	53.2	1.7	1472	3	AAA59241	Aaa59241 Exons D,
	32	53.2	1.7	8056	7	ABZ10246	Abz10246 Haematopo
	33	53	1.7	987	2	AAT90554	Aat90554 Clostridi
c	34	53	1.6	13784	6	ABK40062	Abk40062 Human che
	35	52.4	1.6	3057	2	AAX99555	Aax99555 Nucleic a
	36	52	1.6	972	2	AAX61754	Aax61754 B. burgdo
	37	52	1.6	1029	2	AAX61753	Aax61753 B. burgdo
c	38	52	1.6	5120	4	AAC84677	Aac84677 DNA seque
c	39	52	1.6	110000	2	AAX20248_08	Continuation (9 of
	40	51.8	1.6	8056	7	ABZ10100	Abz10100 Haematopo
	41	51.6	1.6	906	7	ACA28937	Aca28937 Prokaryot
	42	51.6	1.6	2451	4	AAS07045	Aas07045 DNA encod
	43	51.4	1.6	2553	7	ACA52806	Aca52806 Prokaryot
	44	50.8	1.6	2017	9	ADB46163	Adb46163 Plasmodiu
	45	50.8	1.6	5314	9	ADB46161	Adb46161 Plasmodiu

ALIGNMENTS

```
RESULT 1
ACF03365
ID   ACF03365 standard; DNA; 3189 BP.
XX
AC   ACF03365;
XX
DT   11-SEP-2003  (first entry)
XX
DE   Mycoplasma gallisepticum mgc3 gene SEQ ID NO:2.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP1275716-A2.
XX
PD   15-JAN-2003.
XX
PF   11-JUL-2002; 2002EP-00254879.
XX
PR   11-JUL-2001; 2001US-00301572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG ) ZEON CORP.
XX
PI   Okuda T,  Saito S,  Dorsey KM,  Tsuzaki Y;
XX
DR   WPI; 2003-373746/36.
DR   P-PSDB; ABR57375.
XX
PT   DNA molecule derived from a prokaryotic cell, useful for producing a
PT   vaccine for treating viral infections comprises at least one modified DNA
PT   regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS   Claim 9; Page 31-32; 70pp; English.
XX
CC   The present invention describes a DNA molecule derived from a prokaryotic
CC   cell, where at least one of the DNA regions encoding NXB (where N =
CC   asparagine, X = any amino acid other than proline, and B = serine or
CC   threonine) has been modified so that no N-glycosylation occurs during the
CC   expression in a eukaryotic cell. Also described: (1) a fused DNA
CC   molecule, where a DNA encoding a signal sequence has been ligated to the
CC   N-terminal end of the modified DNA molecule as described above so that it
CC   may be expressed as a fusion protein; (2) a recombinant virus integrated
CC   with the DNA molecule or the fused DNA molecule described above; (3)
```

```
CC   producing a modified or fusion protein by using the recombinant virus
CC   described above, to express a protein encoded by the modified DNA
CC   molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC   vaccine comprising the recombinant virus. The DNA molecule has virucide
CC   and immunostimulant activities. The DNA molecule is useful for producing
CC   a vaccine for treating viral infections. The present sequence is used in
CC   the exemplification of the present invention
XX
SQ   Sequence 3189 BP; 1114 A; 576 C; 539 G; 960 T; 0 U; 0 Other;

Query Match       100.0%;   Score 3189;   DB 7;   Length 3189;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 3189;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60

Qy       61 GCTCTTGGTTCTGCGCAAGCTTTGGCTTTAAGCAATCAATACAGAGTAACGATAACACGCAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GCTCTTGGTTCTGCGCAAGCTTTGGCTTTAAGCAATCAATACAGAGTAACGATAACACGCAA 120

Qy      121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTCGTTAGACTTGCAGGTCTTGGACAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTCGTTAGACTTGCAGGTCTTGGACAA 180

Qy      181 AATGGTTCGTTGTTCAATACAGTTCTTAGATGTTGATGATAACTTTATAACAGCAGCT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AATGGTTCGTTGTTCAATACAGTTCTTAGATGTTGATGATAACTTTATAACAGCAGCT 240

Qy      241 AATGGAACAATATCAAATTAGATAGTTTTACTAAAACCATTATATGTTTAGATCTAAGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AATGGAACAATATCAAATTAGATAGTTTTACTAAAACCATTATATGTTTAGATCTAAGT 300

Qy      301 GATGATTGTGTGGATACAAAGTAAAACAAATAGTTTCAGATTACCAACTAGCAGAAAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GATGATTGTGTGGATACAAAGTAAAACAAATAGTTTCAGATTACCAACTAGCAGAAAT 360

Qy      361 AGATTTGATCAAGACAAAGACAAGAAGAGCATATATGCTCCTGTTGGTTAATGATGAAGCTAAC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AGATTTGATCAAGACAAAGACAAGAAGAGCATATATGCTCCTGTTGGTTAATGATGAAGCTAAC 420

Qy      421 GTTCATTTAAAAAGAAGAATTAATACTAACTAACCAAGTGTAATAGAATTGGTAGACAACAATCT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTTCATTTAAAAAGAAGAATTAATACTAACTAACCAAGTGTAATAGAATTGGTAGACAACAATCT 480

Qy      481 AAGTTTGTAATTGGTGTGTGTAATTGGTGTGTGTGATATCCAGCTCACGTAATTAGATTTACTGATGATGGG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AAGTTTGTAATTGGTGTGTGTAATTGGTGTGTGTGATATCCAGCTCACGTAATTAGATTTACTGATGATGGG 540

Qy      541 ACTAAATTTAATTTAATTTACAAACCAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 ACTAAATTTAATTTAATTTACAAACCAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600

Qy      601 GCGCCAATCTTACCTAAAGATTACACCCGAGTTGTTCCTTGGGCAATGGTATATAACTTATACATTCAAAGAAAG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GCGCCAATCTTACCTAAAGATTACACCCGAGTTGTTCCTTGGGCAATGGTATATAACTTATACATTCAAAGAAAG 660

Qy      661 ATCTTACCAAATGACGTCAACACTGAGTGTTGATTGTGGGAATGGTCAAATAACTTATAACTAATACAGATCCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 ATCTTACCAAATGACGTCAACACTGAGTGTTGATTGTGGGAATGGTCAAATAACTTATAACTAATACAGATATT 720

Qy      721 ACAAATGCTGACTAAAACCACTAACTACTGATAATCCTTCAAAATCCTTTAATTCAGGAGCA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ACAAATGCTGACTGATGATGGATGGATGGTCAAATAACTTATAACTAATACAGATATT 780

Qy      781 ATTGCTCAAACTAAAACAATAGATACGATTCTCAATCTGATAATCCTTTAATTCAGGAGCA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 ATTGCTCAAACTAAAACCACTAACCACTACTGATAATCCTTTAATCTTTAATTCAGGAGCA 840

Qy      841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAAACA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAAACA 900
```

```
Qy      901 TCTTTCCAATTAGATGAGAAAAATTGTTGTTATCCAGAATGGACTGGTTCTGAAGAGAATAA 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TCTTTCCAATTAGATGAGAAAAATTGTTGTTATCCAGAATGGACTGGTTCTGAAGAGAATAA 960

Qy      961 AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020

Qy     1021 ATACCCGGACTCCACAAGTTACTTAAAGAAGAGATTCAGTTAACGTATTTTCAAGACTA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 ATACCCGGACTCCACAAGTTACTTAAAGAAGAGATTCAGTTAACGTATTTTCAAGACTA 1080

Qy     1081 TACTTAAACTCAGTAATTCTTATCATTCATTGGTGATAGTATTATATATTTTGGTACC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TACTTAAACTCAGTAATTCTTATCATTCATTGGTGATAGTATTATATATTTTGGTACC 1140

Qy     1141 TCTGAATTACCATCATTATGTACTATTCATTCCCAACTAGATTATCTGATCGATCAACCGCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 TCTGAATTACCATCATTATGTACTATTCATTCCCAACTAGATTATCTGATCGATCAACCGCT 1200

Qy     1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACACA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACACA 1260

Qy     1261 AACCGAACAACGACAACAGCTGATCATCTAGTGGTTCAACAGTGCTGGAACAGGAAAT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 AACCGAACAACGACAACAGCTGATCATCTAGTGGTTCAACAGTGCTGGAACAGGAAAT 1320

Qy     1321 ACTACTAACAACTTCTCAAACAGTTCTAATCCTACTTAAATACTTATCCGTAGTTTTGA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 ACTACTAACAACTTCTCAAACAGTTCTAATCCTACTTAAATACTTATCCGTAGTTTTGA 1380

Qy     1381 ATTGATAGTAAACAACTTCTGCAAACAAAATAGTGAACTAATTGGGCAGATCTAGTTCCTAAC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 ATTGATAGTAAACAACTTCTGCAAACAAAATAGTGAACTAATTGGGCAGATCTAGTTCCTAAC 1440

Qy     1441 GTTATTGAAGCAAGAATATGTCGAATACAGATTAGGTATTCAAAATGAAATTCCAATA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 GTTATTGAAGCAAGAATATGTCGAATACAGATTAGGTATTCAAAATGAAATTCCAATA 1500

Qy     1501 ACTAATGCAGGAACTTTATCCGAAACAACAATGGTGTGTTTTTACTTCAACAGGT 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 ACTAATGCAGGAACTTTATCCGAAACAACAATGGTGTGTTTTTACTTCAACAGGT 1560

Qy     1561 TCAAGAGTAGTTTTAAGAGCTTCTATAACGGTGATCAACGTCCAACTTCCAA 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 TCAAGAGTAGTTTTAAGAGCTTCTATAACGGTGATCAACGTCCAACTTCCAA 1620

Qy     1621 CCTTTCTTATACGTATTTGGTTATTGGTTATTTGGTATTTAGGATAACAACAGGAACTTTCTGG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 CCTTTCTTATACGTATTTGGTTATTGGTTATTTGGTATTTAGGATAACAACAGGAACTTTCTGG 1680

Qy     1681 TACGGAACATATAGGCTTTTAAACAACAGCGCCCTTACGACGTATTAGATTCTCCAAGAGTA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 TACGGAACATATAGGCTTTTAAACAACAGCGCCCTTACGACGTATTAGATTCTCCAAGAGTA 1740

Qy     1741 GGTACTGAAACCAATCAATTTAGGAGAACTTCATTAACTACTCCATATATAAGAGACACAAGGTGGATAT 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 GGTACTGAAACCAATCAATTTAGGAGAACTTCATTAACTACTCCATATATAAGAGACACAAGGTGGATAT 1800

Qy     1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATATCTGGCTATTCTGATAGTAGTACATT 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATATCTGGCTATTCTGATAGTAGTACATT 1860

Qy     1861 ACACCAGAAAGCCGGAAGCATCTTCCAATCGGCTATTCTTTAATTCAGGAGCA 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 ACACCAGAAAGCCGGAAGCATCTTCCAATCGGCTATTCTTTAATTCAGGAGCA 1920

Qy     1921 CAATCAGTTTTAGGATTTGATGGAATTGATGATGGAATTAGAGAATTAAAGCATCA 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 CAATCAGTTTTAGGATTTGATGGAATTGATGATGGAATTAGAGAATTAAAGCATCA 1980
```

```
Qy   1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACACA 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACACA 2040

Qy   2041 TACTTAAGATCACAAATGGATTAGCTAGAACATCTGGATTACCAAACCAACAACCATTC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TACTTAAGATCACAAATGGATTAGCTAGAACATCTGGATTACCAAACCAACAACCATTC 2100

Qy   2101 GGAACAACTCACCAAGTTATTTCAGTATCGATCAGTTCTCATCATCAATTAAGAAT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GGAACAACTCACCAAGTTATTTCAGTATCGATCAGTTCTCATCATCAATTAAGAAT 2160

Qy   2161 ATTAGAACAAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATTAGAACAAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT 2220

Qy   2221 AAATCTAGTGTTTATATACATTAGCTGACTGAGTCAAGTAACCCTGATGCGTCAAGCTCA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AAATCTAGTGTTTATATACATTAGCTGACTGAGTCAAGTAACCCTGATGCGTCAAGCTCA 2280

Qy   2281 TTCAGTCCAACAAGTTTAATTGACGTTAATGAAATTGTGTAATCTTACCTTTATTAGAC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TTCAGTCCAACAAGTTTAATTGACGTTAATGAAATTGTGTAATCTTACCTTTATTAGAC 2340

Qy   2341 AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTCTTCTCATCAAACCCTGGT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTCTTCTCATCAAACCCTGGT 2400

Qy   2401 TCTCCCTGGATCATATACGCTGTGTAAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCTCCCTGGATCATATACGCTGTGTAAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2460

Qy   2461 GAAGGTTCTCGTGCTAAGTATACATCTGATTTCTGGGAACAATCCAATTCAAACCCGAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GAAGGTTCTCGTGCTAAGTATACATCTGATTTCTGGGAACAATCCAATTCAAACCCGAT 2520

Qy   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGCCTAGAAACTTCGTTACAAACCAA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGCCTAGAAACTTCGTTACAAACCAA 2580

Qy   2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640

Qy   2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700

Qy   2701 TACTTAGATGGTAAGTAGTATTATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TACTTAGATGGTAAGTAGTATTATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 2760

Qy   2761 TATAACAACTTTGGCGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TATAACAACTTTGGCGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820

Qy   2821 GGTATTCTTGCAATTATGATCATCTTAGGATCATCTTAGGATTCCTTTAAGAGCTCAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GGTATTCTTGCAATTATGATCATCTTAGGATCATCTTAGGATTCCTTTAAGAGCTCAA 2880

Qy   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940

Qy   2941 GCTGTTGGTTGGTTCAGTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GCTGTTGGTTGGTTCAGTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000

Qy   3001 GCTGCTTTAGGTGCTCGTGTAAATCGTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GCTGCTTTAGGTGCTCGTGTAAATCGTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCT 3060

Qy   3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGCCTAA 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGCCTAAA 3120

Qy   3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTCGCTCCAAAAACCACCGCTCCC 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTCGCTCCAAAAACCAACCGCTCCC 3180

Qy   3181 AAAGAATAA 3189
          |||||||||
Db   3181 AAAGAATAA 3189
```

RESULT 2
AAT75087
ID   AAT75087 standard; DNA; 3189 BP.
XX
AC   AAT75087;
XX
DT   27-FEB-1998 (first entry)
XX
DE   DNA encoding antigenic protein derived from Mycoplasma gallisepticum.
XX
KW   Mycoplasma gallisepticum antigen; epitope; monoclonal antibody; bird;
KW   mycoplasma infection; viral vaccine; avipoxvirus; herpesvirus; ss.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..3189
FT   /*tag= a
FT   /product= "Mycoplasma_antigen"
FT   CDS   1..3189
FT   /*tag= b
FT   /codon= (seq:tga, aa:Trp)
XX
PN   WO9724370-A1.
XX
PD   10-JUL-1997.
XX
PF   27-DEC-1996; 96WO-JP003863.
XX
PR   28-DEC-1995; 95JP-00352754.
XX
PA   (JAPG) NIPPON ZEON KK.
XX
PI   Fujisawa A, Yoshida S;
XX
DR   WPI; 1997-363621/33.
DR   P-PSDB; AAW22721.
XX
PT   Antigenic protein derived from Mycoplasma gallisepticum - useful in
PT   vaccines against, and for diagnosis of mycoplasma infection in birds.
XX
PS   Claim 4; Page 24-31; 45pp; Japanese.
XX
CC   This sequence represents DNA encoding an antigenic protein of Mycoplasma
CC   gallisepticum having an epitope specifically recognised by the monoclonal
CC   antibody (MAb) 35A6. This antigenic protein inhibits the metabolism of
CC   bird infecting mycoplasma. The antigenic protein and the DNA molecule can
CC   be used in the preparation of component and viral vaccines against
CC   mycoplasma infection in birds. Mycoplasma infection can be diagnosed by
CC   hybridising DNA in a sample with a DNA probe corresponding to 10 or more
CC   (preferably 15 to 40) bases of the antigenic protein gene
XX
SQ   Sequence 3189 BP; 1130 A; 572 C; 522 G; 965 T; 0 U; 0 Other;

Query Match   98.6%; Score 3144.2; DB 2; Length 3189;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3161; Conservative   0; Mismatches 28; Indels   0; Gaps   0;

Qy   1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
```

Qy   61 GCTCTTGGTTCTGCGAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCTCTTGGTTCTGCGAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120

Qy 121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGTCTTGGACAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGTCTTGGACAA 180

Qy 181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240

Qy 241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 300

Qy 301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GATGATTGTGCTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 360

Qy 361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTAATGCTAATGAAGCTAAC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTAATGATGAATGAAGCTAAC 420

Qy 421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT 480

Qy 481 AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATATTACTGATGGGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 AAGTTTGTAATTGGTGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGG 540

Qy 541 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600

Qy 601 GCGCCAATCTTACCCTAAAGATTTACACCCAGATTGATTGGTATAACTTATACATTCAAAGAAAG 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GCGCCAATCTTACCCTAAAGATTTACACCCAGATTGATTGGTATAACTTATACATTCAAAGAAAC 660

Qy 661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTTGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGACCAGTAGGTAGAGTTAGTTGA 720

Qy 721 ACAAATGCTAGATGATGGGATGTTGATTGTTGGGAATGGTCAAATAACTAACTATACAGATCCT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 ACAAATGCTGATGATGGGATGTTGATTGTTGGGAATGGTCAAATAACTAACTATACAGATCCT 780

Qy 781 ATTGCTCAAACTAAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGGA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 ATTGCTCAAACTAAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA 840

Qy 841 ATGCCTGGTGCAAACATAGTACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 ATGCCTGGTGCAAACATAGTACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 900

Qy 901 TCTTTCCAATTAGATGAAAAATTGTTTATTCCAGAATGACTCGGTTCTGAAGAGAATAA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 TCTTTCCAATTAGATGAAAAATTGTTTATTCCAGAATGACTCGGTTCTGAAGAGAATAA 960

Qy 961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 AATATTACAAGATCACGAATTGGATGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020

Qy 1021 ATACCCGGGACTCCACAAGTTACTTAAAAGAAGAATTCAGTTAACGTAACGTATTTTCAAGACTA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ATACCAGGGACTCCACAAGTTACTTAAAAGAAGAATTCAGTTAACGTAACGTATTTTCAAGACTA 1080

Qy 1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTCGTGATAGTATTTTATATTTTTGGTACC 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTCGTGATAGTATTTTATATTTTTGGTACT 1140

Qy 1141 TCTGAATTACCATCATCATATGGTACTATTCATTCCCAACTAGAATTATCTGATCTAACCGCT 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TCTGAATTACCATCATCATATGGTACTATTCATTCCCAACTAGAATTATCTGATCTAACCGCT 1200

Qy 1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260

Qy 1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320

Qy 1321 ACTACTAACACTTCTCAAACAGTTTCTCAATCCTAAATACTTAAATACTTAAATCGTAGTTTTGGA 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 ACTACTAACACTTCTCAAACAGTTTCTCAATCCTAAATACTTAAATACTTAAATCGTAGTTTTGGA 1380

Qy 1381 AATTGATAGTAAACCAACTTCTGCAAACAAAATAGTGAAAACTAAATTGGGCAGATCCTAAC 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 AATTGATAGTAAACCAACTTCTGCAAACAAAATAGTGAAAACTAAATTGAGCAGATCCTAAT 1440

Qy 1441 GTTATTGAAGCAAGAATATATGCTGAATACGATAGGTATTCAAAATGAAATTCCAATA 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTTATTGAAGCAAGAATATATGCTGAATACGATAGGTATTCAAAATGAAATTCCAATA 1500

Qy 1501 ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTTACTTCAACAGGT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTTACTTCAACAGGT 1560

Qy 1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTGATCAACGTCCAACTGGAAACTTCCAA 1620

Qy 1621 CCTTTCTTATACGTATTTGGTATTTAGGATCAGTTGTTAATGACTTCATTTTAGAT 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CCTTTCTTATACGTATTTGGTATTTAGGATCAGTTGTTAATGACTTCATTTCTGA 1680

Qy 1681 TACGGAACATATAAGCTTTAAACAACAGCCCTAGCACGTATTAGATTCTCCAAGAGTA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TACGGAACATATAAGCTTTAAACAACAGCCCTAGCACGTATTAGATTCTCCAAGAGTA 1740

Qy 1741 GGTACTGAAAACCAATCAATTAGAAGAACTTCATTAACATACATACCCTGTTGTATGGGTGGATAT 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GGTACTGAAAACCAATCAATTAGAAGAACTTCATTAACATACATACCCTGTTATGGGTGGATAT 1800

Qy 1801 CTAACTGAAGAAGCTGCTAGAAGTTTCTCTAATACTCCATATAAGAGCACAAGGTGAC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATAAGAGCACAAGGTGAC 1860

Qy 1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAAATACTTATGAGTACATT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT 1920

Qy 1921 CAATCAGTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 CAATCAGTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA 1980

Qy 1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATAGAGAATAAT 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATAGAGAATAAT 2040

Qy 2041 TACTTAAGATCACAAATTGGATTAGCTGAATTACCAACCACCATTC 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 TACTTAAGATCACAAATTGGATTAGCTGAATTACCAACCACCATTC 2100

Qy 2101 GGAACAACTCACCAAGTTATTTCAGTATCACCTGGTATCAGTTCTCATCAATTAAGAAT 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 GGAACAACTCACCAAGTTATTTCAGTATCACCTGGTATCAGTTCTCATCAATTAAGAAT 2160

Qy 2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATCGTATGGTACTTCTTATTCACAAATGAAAATAAT 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCACAAATGAAAATAAT 2220

Qy 2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280

XX
DT 11-AUG-2003 (first entry)
XX
DE DNA encoding Mycoplasma gallisepticum live vaccine.
XX
KW Vaccine; ds; vaccination; cytadherence-deficiency; virucide.
XX
OS Mycoplasma gallisepticum.
XX
PN US2002187162-A1.
XX
PD 12-DEC-2002.
XX
PF 19-APR-2002; 2002US-00125818.
XX
PR 21-APR-2001; 2001US-0285569P.
XX
PA (GEAR/) GEARY S J.
PA (SILB/) SILBART L.
PA (MARC/) MARCUS P.
PA (SEKE/) SEKELLICK M.
XX
PI Geary SJ, Silbart L, Marcus P, Sekellick M;
XX
DR WPI; 2003-341017/32.
XX
PT New method of avian vaccination against virulent strains of Mycoplasma
PT gallisepticum by administering to a bird an immunogen comprising a
PT cytadherence-deficient M. gallisepticum.
XX
PS Claim 21; Page 6-10; 12pp; English.
XX
CC This invention relates to a novel avian vaccination against virulent
CC strains of Mycoplasma gallisepticum. The vaccination comprises
CC administering to a bird an immunogen comprising a cytadherence-deficient
CC M. gallisepticum having an inability to express at least two of three of
CC the following proteins expressed by wild-type M. gallisepticum;
CC cytadhesin molecule GapA; crmA protein; 45 kDa protein. The composition
CC of the invention may have virucide activity and may be used as a
CC virucide. The vaccine is useful for preventing M. gallisepticum virus
CC infection in birds. The present sequence represents the DNA sequence
CC encoding the Mycoplasma gallisepticum vaccine of the invention. This
CC sequence contains the coding sequences of the Gap-A, crmA and 45kDa
CC proteins which are not expressed in cytadherence deficient M.
CC gallisepticum
XX
SQ Sequence 8354 BP; 2916 A; 1463 C; 1426 G; 2545 T; 0 U; 4 Other;

Query Match   98.6%; Score 3143.4; DB 7; Length 8354;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3159; Conservative 2; Mismatches 28; Indels 0; Gaps 0;

```
Qy         1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3778 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 3837

Qy        61 GCTCCTTGGTCTTGCAAGCTTTGGCTTTAAGCAATCAATCAGATAAGAGTAACGAATAACACGCAA 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3838 GCTCCTTGGTCTTGCAAGCTTTGGCTTTAAGCAATCAATCAGATAAGAGTAACGAATAACACGCAA 3897

Qy       121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTGTTAGACTTGCAGGTCTTGGACAA 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3898 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTGTTAGACTTGCAGGTCTTGGACAA 3957

Qy       181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3958 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 4017

Qy       241 AATGGAACAATATCAAATTAGAGATAGTTTTACTAAACCATTATATGGTTAGATCTAAGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4018 AATGGAACAATATCAAATTAGAGATAGTTTTACTAAACCATTATATGGTTAGATCTAAGT 4077

Qy       301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
```

```
Db  5158  ATTGATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATGAGCAGATCCTAAT  5217
                                                                      ||||||||
Qy  1441  GTTTATTGAAGCAAGAATATATGCTGAATACAGATTAGTAGTTATTCAAAATGAAATTCCAATA  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5218  GTTTATTGAAGCAAGAATATATGCTGAATACAGATTAGTAGTTATTCAAAATGAAATTCCAATA  5277

Qy  1501  ACTAATGCAGGAAACTTATCCGAAACACACAATTGGTGGTGGTTGGTTTTACTTCAACAGGT  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5278  ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGGTTGGTTTTACTTCAACAGGT  5337

Qy  1561  TCAAGAGTAGTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5338  TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA  5397

Qy  1621  CCTTTCTTATACGTATTTGGTTATTTGGTTATTTAGGATACCAACAAACTAGACAGGAACTTTCTGG  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5398  CCTTTCTTATACGTATTTGGTTATTTGGTTATTTAGGATACCAACAAACTAGACAGGAACTTTCTGA  5457

Qy  1681  TACGGACATATAAGCTTTTAAACAACAGCGCCCTTACGACGTATTAGATTCTCCAAGAGTA  1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5458  TACGGACATATAAACTATTAAACAACAGCCCTTACGACGTATTAGATGCTGCAAGAGTA  5517

Qy  1741  GGTACTGAAAACCAATCAATTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATAT  1800
                                                                     :
Db  5518  GGTACTGAAAACCAATCAATTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATAY  5577

Qy  1801  CTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATAAGAGCACAAGGTGAC  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5578  CTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATAAGAGCACAAGGTGAC  5637

Qy  1861  ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT  1920
                                                             :
Db  5638  ACACCAGAAAGCCGAAGCATCTTCCAATCTGCTAYTCTGATAATACTTATGAGTACATT  5697

Qy  1921  CAATCAGTTTAGGATTTGATGAATTAGAAATAACTAAAATGTTGGGGTAAAGCATCA  1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5698  CAATCAGTTTAGGATTTGATGAATTAGAAATAACTAAAATGTTGGGGTAAAGCATCA  5757

Qy  1981  AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTCGAGAAATGATTGCTCGAACAACA  2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5758  AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTCGAGAAATGATTGCTCGAACAACA  5817

Qy  2041  TACTTAAGATCACAAATTGGATTAGCTTAGAACATCTCGGATTACCAACCAACAACCATTC  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5818  TACTTAAGATCACAAATTGGATTAGCTTAGAACATCTCGGATTACCAACCAACAACCATTC  5877

Qy  2101  GGAACAACTCACCAAGTTATTCAGTATGATCAGCTGATCACGTCTCATCAATTAGAAT  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5878  GGAACAACTCACCAAGTTATTCAGTATGATCAGCTGATCACGTCTCATCAATTAGAAT  5937

Qy  2161  ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAAATGAAAATAAT  2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5938  ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAAATGAAAATAAT  5997

Qy  2221  AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGTCA  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5998  AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGTCA  6057

Qy  2281  TTCAGTCCAACAGTTTAATTGACGTTAATGAAATTGGTAATCTTTATTAGAC  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6058  TTCAGTCCAACAGTTTAATTGACGTTAATGAAATTGGTAATCTTTATTAGAC  6117

Qy  2341  AATTCAATTCTATACAGTAAATGCTCGTCGTAATGTTGCATTGTTCTCATCAACCCTGGT  2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6118  AATTCAATTCTATACAGTAAATGCTCGTCGTAAATGTTGCATTGTTCTCATCAAACCCTGGT  6177

Qy  2401  TCTCCTGGATCATATACTGCTGTAAATCCTCGTAAATACATTCGATATTGCTTTT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6178  TCTCCTGGATCATATACTGCTGTAAATCCTCGTAAATACATTCGATATTGCTTTT  6237

Qy  2461  GAAGGTTCGGTGCTAAGTAACTAATTGGGCAGATCCTAAC  2520
          |||||||||||||||||||||||||||||||||||||||||
Db  6238  GAAGGTTCGGTGCTAAAATATACATTCGATTTCTGAGGAACAATTCAAACCCGAT  6297
```

```
Db  4078  GATGATTTTGCTGGATACAAGTAAAACAAATAGTTTCAGATTACCAACTAGTGATGAAGCTAAC  4137
                                                           |||||||||||||||||
Qy  361   AGATTTGATCAAGCAAGAGCAAAACAAGAGCATATATGCTCTGTTGGTTAATGAAGCTAAC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4138  AGATTTGATCAAGCAAGACAAAACAAGAGCATATATGCTCTGTTGGTTAATGAAGCTAAC  4197

Qy  421   GTTCATTTAAAAAGAATTAATACTAACTAACTAATGGTAATAGAAGAACAACAATTCT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4198  GTTCATTTAAAAAGAATTAATACTAACTAACTAATGGTAATAGAAGAACAACAATTCT  4257

Qy  481   AAGTTGTATTGGTGGTGTTGATAATCCAGCTCACGTAATTACTACTGATGATGGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4258  AAGTTGTATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTACTGATGATGGG  4317

Qy  541   ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATGTTAATGACTTCATTTTAGAT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4318  ACTAAATTTAATTTTACAAAGCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  4377

Qy  601   GCGCCAAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTCATTCAAAGAAAG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4378  GCGCCAAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  4437

Qy  661   ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4438  ATCTTACCAAATGACGTAAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA  4497

Qy  721   ACAAATGCTGATGAGTGGGATGTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4498  ACAAATGCTGATGAGTGGGATGTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT  4557

Qy  781   ATTGCTCAAACTAAAACCACTACTGATAATCAAATCCTTCAACTTTAATTCAGGAGACA  840
                                                    :
Db  4558  ATTGCTCAAACTAAAACCACTACTGATAATCCTTCAACTTTAATTCAGGAGACA  4617

Qy  841   ATGCCCGGTGCAAACAACAATAGATAACGATTCTCAATTGAAATGTCAAGCATAGAAATAAAACA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4618  ATGCCCGGTGCAAACAATAGATAACGATTCTCAATTGAAATGTCAAGCATAGAAATTAAAAACA  4677

Qy  901   TCTTTCCAATTAGATGAAAAATTTGTTTATCCAGAATGACTGGTTCTGAAGAGAATAAA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4678  TCTTTCCAATTAGATGAAAAATTTGTTTATCCAGAATGACTGGTTCTGAAGAGAATAAA  4737

Qy  961   AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAACGAAAGATATTGGATTCTTGAC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4738  AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAACGAAAGATATTGGATTCTTGAC  4797

Qy  1021  ATACCCGGGACTCCACAAGTTACTTTAAAGAAGATTCAGTTAAGTATTTTCAAGACTA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4798  ATACCACGGGACTCCACAAGTTACTTTAAAGAAGATTCAGTTAAGTATTTTCAAGACTA  4857

Qy  1081  TACTTAAACTCAGTTAATTCTTTACCATTCATTGTGATGATTTTATATTTTTGGTACC  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4858  TACTTAAACTCAGTTAATTCTTTATCATTCATTGTGATGATTTTATATTTTTTGGTACT  4917

Qy  1141  TCTGAATTACCATCATTATGTGTACTATTCATTCATTCCCAACTAGATTATCTGATCTAACCGCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4918  TCTGAATTACCATCATTATGTGTACTATTCATTCATTCCCAACTAGATTATCTGATCTAACCGCT  4977

Qy  1201  TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACCGTACAACAACA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4978  TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACCGTACAACAACA  5037

Qy  1261  AACGGAACAACGACAACAGAGCTTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5038  AACGGAACAACGACAACAGAGCTTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  5097

Qy  1321  ACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTAAATACTTAGTTTTGGA  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5098  ACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTTAAATACTTAGTTTTGGA  5157

Qy  1381  ATTGATAGTAAACCAACTTCTGCAAACAAATAGATGAAACTAATGGGCAGATCCTAAC  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy    2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTCGGCTAGAAACTTCGTTACAAACCAA 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6298 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTCGGCTAGAAACTTCGTTACAAACCAA 6357

Qy    2581 AGCTTCCTTAAACAGTTTAGTTGACTTGACTTCACTCCTGCTAATGCTGGTACTAACCGTGTA 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6358 AGCTTCCTTAAACAGTTTAGTTGACTTGACTTCACTCCTGCTAATGCTGGTACTAACCGTGTA 6417

Qy    2641 GTGGTTGATCCGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6418 GTGGTTGATCCGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 6477

Qy    2701 TACTAGATGGTAAGTATTATGATGCTAAATTAAGAACAATAATTAGTAACATTCTCT 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6478 TACTAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 6537

Qy    2761 TATAACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6538 TATAACAACTTTCGCTGCGTTACCTTCATGAGTAGTGCCTACAGCAATTGGTAGTACATTA 6597

Qy    2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATTCGGTATTCCTTTAAGAGCTCAA 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6598 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATTCGGTATTCCTTTAAGAGCTCAA 6657

Qy    2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6658 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 6717

Qy    2941 GCTGTTGGTTCAGTTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCT 3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6718 GCTGTTGGTTCAGTTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCT 6777

Qy    3001 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGCTCTAAACCTGCT 3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6778 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGCTCTAAACCTGCT 6837

Qy    3061 GCTCCAGCTAAACCATTCGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGCCTAAA 3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6838 GCTCCAGCTAAACCATTCGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGCCTAAA 6897

Qy    3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6898 TCTGGTGCGCCTACAAAACCAACTGCTCCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 6957

Qy    3181 AAGAATAA 3189
           ||||||||
Db    6958 AAGAATAA 6966
```

RESULT 4
AAF88025
ID AAF88025 standard; DNA; 10651 BP.
XX
AC AAF88025;
XX
DT 16-JUL-2001 (first entry)
XX
DE Vector pcDNA3-FasL-IRES-crmA DNA.
XX
KW Vector; pcDNA3-FasL-IRES-crmA; autoimmune disease; apoptosis; crmA;
KW suicide enzyme; antiinflammatory; immunomodulatory; rheumatic; arthritic;
KW dermatological; gene therapy; rheumatoid arthritis;
KW systemic lupus erythematosus; Sjoegren-Syndrome; polymyositis;
KW dermatomyositis; polymyalgica; rheumatism; arteritis temporalis;
KW Crohn's disease; Bechterew disease; colitis ulcerosa; thyroiditis;
KW autoimmune hepatitis; diabetes mellitus (type I); psoriasis; dermatitis;
KW pemphigus vulgaris; multiple sclerosis; myasthenia gravis;
KW viral infection; brain inflammation; transplant rejection; ds.
XX
OS Synthetic.
XX
PN WO200127254-A2.

XX
CC This invention describes a novel gene transfer vector (N1) which
CC comprises (1) a first nucleic acid sequence (I) encoding one or more
CC ligands that trigger apoptosis; (2) a second nucleic acid sequence (II)
CC encoding one or more antigens; and optionally (3) a third nucleic acid
CC sequence (III) encoding one or more anti-apoptosis molecules; and
CC optionally (4) a fourth nucleic acid sequence (IV) encoding one or more
CC suicide enzymes. The products of the invention have antiinflammatory,
CC immunomodulatory, rheumatic, arthritic and dermatological activity and
CC can be used in gene therapy. The gene vector may be used to produce
CC medicine for the prevention or therapy of autoimmune diseases, such as
CC rheumatoid arthritis, systemic lupus erythematosus, Sjoegren-Syndrome,
CC polymyositis, dermatomyositis, polymyalgica, rheumatism, arteritis
CC temporalis, Crohn's disease, Bechterew disease, colitis ulcerosa,
CC autoimmune hepatitis, thyroiditis, diabetes mellitus (type I), suprarena body
CC insufficiency, thyroiditis, psoriasis, dermatitis, pemphigus vulgaris,
CC multiple sclerosis, myasthenia gravis, or for the prevention or therapy
CC of diseases associated with chronic inflammation, that target
CC immunopathogens, particulary inflammations associated with viral or
CC bacterial infection, particularly hepatitis B or C infections or of the
CC inflammation of the brain after infection from Masern virus, or
CC transplant rejections. The vector may be used in ex vivo modification of
CC animal or mammal cells, particularly human cells. This sequence
CC represents the vector pcDNA3-FasL-IRES-crmA which contains the suicide
CC enzyme crmA
XX
SQ Sequence 10651 BP; 2782 A; 2618 C; 2490 G; 2759 T; 0 U; 2 Other;

Query Match   98.6%; Score 3143.4; DB 5; Length 10651;
Best Local Similarity   99.1%; Pred. No. 0;
Matches 3159;   Mismatches 28;   Conservative 2;   Indels 0;   Gaps 0;

```
Qy      1 ATGAATAATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6 ATGAATAATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA  65

Qy     61 GCTCCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGCAATCAGAGATAACGATAACACGCAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 GCTCCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGCAATCAGAGATAACGATAACACGCAA 125

Qy    121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTCGTTAGACTTCGGTCTTGGACAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTCGTTAGACTTCGGTCTTGGACAA 185

Qy    181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 245

Qy    241 AATGGAACAATTATCAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 AATGGAACAATTATCAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 305

Qy    301 GATGATTGGTGGTGGATCAAAGTAAAAGTAAAACAAAGAGTTTCAGATTACACAACTAGCAGAAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 GATGATTTGTGGTGGATCAAAGTAAAAGTAAAACAAAGAGTTTCAGATTACACAACTAGCAGAAAT 365
```

```
Qy  1446 GTTATTGAAGCAAGAATATATGTCGAATACGACAATTGGTGTGTGGTTGGTTTACTTCAACAGGTA 1505
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       

Qy  1501 ACTAATGCAGGAAAACTTATCGAAACACAATTGGTGGTGTGGTTGGTTTACTTCAACAGGT 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1506 ACTAATGCAGGAAAACTTTATCCGAAACACAATTGGTGGTGTGGTTGGTTTACTTCAACAGGT 1565

Qy  1561 TCAAGAGTAGTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1566 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA 1625

Qy  1621 CCTTTCTATACGTATTGGTTATTGGTTAGGATACCAACAAACATAGAACAGGAACTTTCTGG 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1626 CCTTTCTATACGTATTGGTTATTGGTTAGGATCCAACAAACATAGAACAGGAACTTTCTGA 1685

Qy  1681 TACGGAACATATAAGCTTTAAACAACAACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1686 TACGGAACATATAAACATTAAACAACAACAGCCCTTACGACGTATTAGATGCTGCAAGAGTA 1745

Qy  1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACTAACATCCATATAGAGACAAGGTGAC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db  1746 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACTAACATCCATATAGAGACAAGGTGAC 1805

Qy  1801 CTAACTGAAGAAGCCGAAGCATCTTCCAATCTGCCTATTCTGATAATACTTATGAGTAGTACATT 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1806 CTAACTGAAGAAGCCGAAGCATCTTCCAATCTGGCTAYTCTGATAATACTTATGAGTAGTACATT 1865

Qy  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATAACTTAAATGTTGGGGTTAAAGCATCA 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1866 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTAYTCTGATAATAACTTAAATGTTGGGGTTAAAGCATCA 1925

Qy  1921 CAATCAGTTTAGGATTTGATGGAATAGAAATAAACTTAAATGTTGGGGTTAAAGCATCA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1926 CAATCAGTTTAGGATTTGATGGAATAGAAATAAACTTAAATGTTGGGGTTAAAGCATCA 1985

Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAATCCTCGAGAAATGATTGCTGCAACAACA 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1986 AGCTTCTTAAACTCAAATAGACCAAATCCAAATCCTCGAGAAATGATTGCTGCAACAACA 2045

Qy  2041 TACTTAAGATCACAAATTGGATTAGCTAGATATTGGATTCTTGAC 2100
         |||||||||||||||||||||||||||||||||||||||||||||
Db  2046 TACTTAAGATCACAAATTGGATTAGCTAGATATTGGATTCTTGAC 2105

Qy  2101 GGAACAACTCACCCAGTTATTTCAGTATCGATCAGTTCTCATCATCAATTAAGAAT 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2106 GGAACAACTCACCCAGTTATTTCAGTATCGATCAGTTCTCATCATCAATTAAGAAT 2165

Qy  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2166 ATTAGAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCACAAATGAAAATAAT 2225

Qy  2221 AAATCTAGTGTTTATACATTAAGATTAGCTAGATTGAAATTGGTGTAATCTTATTAGAC 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2226 AAATCTAGTGTTTATACATTAAGATTAGCTAGATTGAAATTGGTGTAATCTTATTAGAC 2285

Qy  2281 TTCAGTCCAACAGTTAATTGACGTTAATGCGTGTAAATACATTTAATCAGAAACTTATCTGATATTGCTTTT 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2286 TTCAGTCCAACAGTTAATTGACGTTAATGCGTGTAAATACATTTAATCAGAAACTTATCTGATATTGCTTTT 2345

Qy  2341 AATTCATTCTATACGTTGCTCGTGTAAATGGTGTAATTGTTGCATTGTTCTCATCAAACCCTGGT 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2346 AATTCATTCTATACGTTGCTCGTGTAAATGGTGTAATTGTTGCATTGTTCTCATCAAACCCTGGT 2405

Qy  2401 TCTCCCTGGATCATATACTTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2406 TCTCCCTGGATCATATACTTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2465

Qy  2461 GAAGGTTCTGGTGTCTAAGTATACATCTAAGTTCTGGGAACAATCCAATTCAAACCCGAT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2466 GAAGGTTCTGGTGTCTAAATATCATCTAAGTTCTGAGGAACAATCCAATTCAAACCCGAT 2525

Qy  2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   361 AGATTTGATCAAAGACAAGAGCATATTATGTCTCGTCTGTTAATGATGAAGCTAAC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366 AGATTTGATCAAAGACAAGACAAGAGCATATTATGTCTCGTCTGTTAATGATGAAGCTAAC 425

Qy   421 GTTCATTTAAAAAGAATTAATACTAACTAACTAACTGAATAGAAATTGGTAATGAAACAACAATTCT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426 GTTCATTTAAAAAGAATTAATACTAACTAACTAACTGAATAGAAATTGGTAATGAAACAACAATTCT 485

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCACGTAATTAGATTACTGATGATGGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486 AAGTTTGTAATTGGTGGTGTTGATAATCACGTAATTAGATTACTGATGATGGG 545

Qy   541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAAATTGTTAATGACTTCATTTAGAT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546 ACTAAATTTAATTTACAAAGCAAACTCAAGGTGAAAATTGTTAATGACTTCATTTAGAT 605

Qy   601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG 665

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTGATGTGGGAATGGTCAAATAACTAATAACAGATCCT 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGAATGGTCAAATAACTAATAACAGATCCT 725

Qy   721 ACAAATGCTGATGAATAAAACCACTACTGATAATCCTTCAACTTTAATTCAGGAGGCA 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 ACAAATGCTGATGAATGGTCAAATGGTCAAATCCTCAACTTTTAATTCAGGAGGCA 785

Qy   781 ATTGCTCAAACTAAAACCACTACTGATACGGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 ATTGCTCAAACTAAAACCACTACTGATACGGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 845

Qy   841 ATGCCTGGTGCGAAACAATAGATACGGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 ATGCCTGGTGCGAAACAATAGATACGGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 905

Qy   901 TCTTTTCCAATTAGATGAAAAAATTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAA 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 TCTTTTCCAATTAGATGAAAAAATTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAA 965

Qy   961 AATATTACAGAGTTAGCTACTGGGAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 AATATTACAGAGTTAGCTACTGGGAGTTTGCCAAGCAACGAAAGATATTGAATTCTTGAC 1025

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAATTTTCAAGACTA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1026 ATACCAGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAATTTTCAAGACTA 1085

Qy  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTAGTATTTTGGTACC 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1086 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTAGTATTTTTGGTACT 1145

Qy  1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATTCTGATCTAACCGCT 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1146 TCTGAATTACCATCATTATGATGTACATTCATTCCCAACTAGATTATCTGATCTAACCGCT 1205

Qy  1201 TTGAATTCAAGTTAAAACAGGTGTCTGAAGCTTCAACAGGTGCTGAACGTGCTACAACAACA 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1206 TTGAATTCAAGTTAAAACAGACACAACTGATCAATCAGCAGGTGCTGAACGTGCTACAACAACA 1265

Qy  1261 AACGGAACAACGACAGCACAGCAACTGATATTGAAGCTTCAACAGGTGCTGAACAGGAAAT 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1266 AACGGAACAACGACAGCACAGCAACTGATA CATCGAACAGGTGCTGAACAGGAAAT 1325

Qy  1321 ACTACTAACACTTCTCAAACAGTTTCTCAAACAGTTTCTCTACTTAAATCCTACTCGTAGTTTTGGA 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1326 ACTACTAACACTTCTCAAACAGTTTCTCAAACAGTTTCTCTACTTAAATCCTACTCGTAGTTTTGGA 1385

Qy  1381 ATTGATAGTAACAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAAC 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1386 ATTGATAGTAACAACTTCTGCAAACAAAATAGATGAAACTAATTGAGCAGATCCTAAT 1445
```

PI　Okuda T, Saito S, Dorsey KM, Tsuzaki Y;
XX
DR　WPI; 2003-373746/36.
DR　P-PSDB; ABR57376.
XX
PT　DNA molecule derived from a prokaryotic cell, useful for producing a
PT　vaccine for treating viral infections comprises at least one modified DNA
PT　regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS　Example 2; Page 62-64; 70pp; English.
XX
CC　The present invention describes a DNA molecule derived from a prokaryotic
CC　cell, where at least one of the DNA regions encoding NXB (where N =
CC　asparagine, X = any amino acid other than proline, and B = serine or
CC　threonine) has been modified so that no N-glycosylation occurs during the
CC　expression in a eukaryotic cell. Also described: (1) a fused DNA
CC　molecule, where a DNA encoding a signal sequence has been ligated to the
CC　N-terminal end of the modified DNA molecule as described above so that it
CC　may be expressed as a fusion protein; (2) a recombinant virus integrated
CC　with the DNA molecule or the fused DNA molecule described above; (3)
CC　producing a modified or fusion protein by using the recombinant virus
CC　described above, to express a protein encoded by the modified DNA
CC　molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC　vaccine comprising the recombinant virus. The DNA molecule has virucide
CC　and immunostimulant activities. The DNA molecule is useful for producing
CC　a vaccine for treating viral infections. The present sequence is used in
CC　the exemplification of the present invention
XX
SQ　Sequence 3129 BP; 1078 A; 581 C; 541 G; 929 T; 0 U; 0 Other;

Query Match　　95.9%;　Score 3057.8;　DB 7;　Length 3129;
Best Local Similarity　98.7%;　Pred. No. 0;
Matches 3083;　Conservative　0;　Mismatches　42;　Indels　0;　Gaps　0;

Qy　　65　TTGGTTCGCAAGCTTTGGCTTTAAGCAATCAGATAGAGTAACGATAACACGCAATTAG　124
　　　　　 ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db　　 5　TCGGTTCGCAAGCTTTGGCTTTAAGCAATCAGATAGAGTAACGATAACACGCAATTAG　64

Qy　125　TTAATCAAGACAAGAAGCTAGATGCTAATTCTCGTTGCAGTGCTTGGTTAATGATGAAGCAACAAATG　184
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　65　TTAATCAAGACAAGAACCTAGATGCTAATTCTCGTTGCAGTGCTTGGTTAATGATGATGAAGCAACAGG　124

Qy　185　GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATGATGAACTTATAACAGCAGCTAATG　244
　　　　　 |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db　125　GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATGTGATGATGAACTTATAACAGCAGCTCAGG　184

Qy　245　GAACAATTATCAAATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGTGATG　304
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　185　GAACAATTATCAAATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGTGATG　244

Qy　305　ATTGTGGTGGATCACAAGTAAAACAAATAGTTTCAGATTACAACTAGCAGAAATAGAT　364
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　245　ATTTTGCTGGATCACAAGTAAAACAAATAGTTTCAGATTACAACTAGCAGAAATAGAT　304

Qy　365　TTGATCAAGACAAACAGACGAGCTATATTATGCTCGTTGGTTAATGAAGCTAACGTTC　424
　　　　　 |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db　305　TTGATCAAGACAAACAGACGAGCTATATTATGCTCGTTGGTTAATGAAGCTAACGTTC　364

Qy　425　ATTTAAAAAGAATTAATTACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGT　484
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　365　ATTTAAAAAGAATTAATTACTAACTCAAATAGAATTGGTAATAGAATAGAAACCAAATTCTAAGT　424

Qy　485　TTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGGACTA　544
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　425　TTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGGACTA　484

Qy　545　AATTTAATTTACAAACCAAACTCAAGTGAAATTGTTAATGACTTCATTTTAGATGCGC　604
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　485　AATTTCAGTTACACAGCAAACTCAAGTGAAATTGTTAATGACTTCATTTTAGATGCGC　544

Qy　605　CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCT　664
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　545　CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCT　604

Db　2526　GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTCAGTGGCTAGAAACTTCGTTACAAACCAA　2585
Qy　2581　AGCTTCTTAAACAGTTAGTTAGTTGACTTCACTCCTGCTAAATGCTGGTACTAACTACCGTGTA　2640
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2586　AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAAATGCTGGTACTAACTACCGTGTA　2645
Qy　2641　GTGGTTGATCCTGATGGTAATTAACAAACCAAAACCTACCTTCTAAAAGTTCAGATCCAA　2700
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2646　GTGGTTGATCCTGATGGTAATTAACAAACCAAAACCTACCTTCTAAAAGTTCAGATCCAA　2705
Qy　2701　TACTTAGATGGTAAGTATTATGATGCTAAATTAAGAACAATAATTAGTAACATTCTCT　2760
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2706　TACTTAGATGGTAAGTATTATGATGCTAAATTAAGAACAATAATTAGTAACATTCTCT　2765
Qy　2761　TATAACAACTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA　2820
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2766　TATAACAACTTGTGCTTTACCTTCATGAGTAGTGCCTACAGCAATTGGTAGTACATTA　2825
Qy　2821　GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAGAGCTCAA　2880
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2826　GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAGAGCTCAA　2885
Qy　2881　AGAAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAAAGTTGATACCTTGACTGCT　2940
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2886　AGAAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAAAGTTGATACCTTGACTGCT　2945
Qy　2941　GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT　3000
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2946　GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAGAAAAAACCT　3005
Qy　3001　GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGTGCTAAACCTGCT　3060
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3006　GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGTGCTAAACCTGCT　3065
Qy　3061　GCTCCAGCTAAACCATCTGCACCAAAAGTAGCTGCACCAAAAGTAGCTCACCAGCTAAACCAACCTGGGCCTAAA　3120
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3066　GCTCCAGCTAAACCATCTGCACCAAAAGTAGCTGCACCAAAAGTAGCTCACCAGCTAAACCAACCTGCGCGCCTAAA　3125
Qy　3121　TCTGGTGCGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAACCAACCGCTCCC　3180
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3126　TCTGGTGCGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAACCAACCGCTCCC　3185
Qy　3181　AAGAATAA　3189
　　　　　 ||||||||
Db　3186　AAGAATAA　3194

RESULT 5
ACF03435
ID　ACF03435 standard; DNA; 3129 BP.
XX
AC　ACF03435;
XX
DT　11-SEP-2003　(first entry)
XX
DE　Mycoplasma gallisepticum modified mgc3 gene SEQ ID NO:79.
XX
KW　DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW　immunostimulant; viral infection; gene; ds.
XX
OS　Mycoplasma gallisepticum.
XX
PN　EP1275716-A2.
XX
PD　15-JAN-2003.
XX
PF　11-JUL-2002; 2002EP-00254879.
XX
PR　11-JUL-2001; 2001US-00901572.
PR　25-APR-2002; 2002US-00131591.
XX
PA　(JAPG) ZEON CORP.
XX

This page consists of two columns of DNA sequence alignment blocks (Query "Qy" vs Database "Db"), with flanking residue-position numbers.

Right column (positions 1745–2884):

```
Qy  1745 CTGAAACCAATCAATTAGAGAACTTCATTAACAATACCCTGTTATGGGTGGATATCTAA 1804
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1685 CTGAAACCAATCAATTAGAGAACTTCATTAACAATACCCTGTTATGGGTGGATATCTAA 1744

Qy  1805 CTGAAGAAGGTGCTAGAAGTTCTTCTAATACTCCATATATAAGAGCACAAGGTGACACAC 1864
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1745 CTGAAGAAGGTGCTAGAAGTTCTTCTAATACTCCATATATAAGAGCACAAGGTGACACAC 1804

Qy  1865 CAGAAAGCCGGAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATTCAAT 1924
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1805 CAGAAAGCCGGAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATTCAAT 1864

Qy  1925 CAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCAAGCT 1984
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1865 CAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCAAGCT 1924

Qy  1985 TCTTAAACTCAAATAGACCAAATCCAAATCGTTCTAGAAATGATTGCTGCAACAACATACT 2044
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1925 TCTTAAACTCAAATAGACCAAATCCAAATCGTTCTAGAAATGATTGCTGCAACAACATACT 1984

Qy  2045 TAAGATCACAAATTGGATTAGCTAGAAGATATTGGATTCTTGATATCACAACAACCATTCGGAA 2104
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1985 TAAGATCACAAATTGGATTAGCTAGAAGATATTGGATTCTTGATATCACAACAACCATTCGGAA 2044

Qy  2105 CAACTCACCAAGTTATTTCAGTATCACTGGTGATCCAGTTCTCATCAATTAAGAAATATTA 2164
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2045 CAACTCACCAAGTTATTTCAGTATCACTGGTGATCCAGTTCTCATCAATTAAGAAATATTA 2104

Qy  2165 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAATAAAT 2224
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2105 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATCAGAAAT 2164

Qy  2225 CTAGTGTTTATACATTAAGATTAAGCTTAGAGTAACCCTGGTCAAGTAACCCTCATTCA 2284
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2165 CTAGTGTTTATACATTAAGATTAAGCTTAGAGTAACCCTGGTCAAGTAACCCTCATTCA 2224

Qy  2285 GTCCAACAAGTTTAATTGACGTTAATGAAATTGGTAATCTTACCTTTATTGAGACAATT 2344
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2225 GTCCAACAAGTTTAATTGACGTTAATGAAATTGGTAATCTTACCTTTATTGAGACAATT 2284

Qy  2345 CATTCTATACGTAAATGCTGGTGGTAATAGTTGGATTGTTCATGTTCTCATCAATCCTGGTTCTC 2404
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2285 CATTCTATACGTAAATGCTGGTGGTAATAGTTGGATTGTTCATGTTCTCATCAATCCTGGTTCTC 2344

Qy  2405 CTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCGTAATTGCTTTGAAG 2464
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2345 CTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCGTAATTGCTTTGAAG 2404

Qy  2465 GTTCTGGTGCTAAGTATACATCTGATTTCTGGGGACAATCCAATTCAAACCCGATGAGT 2524
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2405 GTTCTGGTGCTAAGTATACATCTGATTTCTGGGGACAATCCAATTCAAACCCGATGAGT 2464

Qy  2525 ACTTAATTCAAAATGGGTTCACTAGTCAAGTCGCTAGAAACTTCGTTACAAACCAAAGCT 2584
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2465 ACTTAATTCAAAATGGGTTCACTAGTCAAGTCGCTAGAAACTTCGTTACACAGCAAAGCT 2524

Qy  2585 TCTTAAACAGTTTAGTTGACTTCACTCCCTGCTAATGCTGGTACTAACCTACCCGTAGTGG 2644
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2525 TCTTAAACAGTTTAGTTGACTTCACTCCCTGCTAATGCTGGTACTAACCTACCCGTAGTGG 2584

Qy  2645 TTGATCCTGATGGTAATTAACAACAAAACCTACCTCTAAAAGTTCAGATCCAATACT 2704
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2585 TTGATCCTGATGGTAATTAACAACAAAACCTACCTCTAAAAGTTCAGATCCAATACT 2644

Qy  2705 TAGATGGTAAGTATTATGATGGTAAATTAAAGAACAATAATTAGTAACATTCTTTATA 2764
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2645 TAGATGGTAAGTATTATGATGGTAAATTAAAGAACAATAATTAGTAACATTCTTTATA 2704

Qy  2765 ACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTAGTAGGTA 2824
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2705 ACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTAGTAGGTA 2764

Qy  2825 TTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTAAGAGCTCAAAGAA 2884
```

Left column (positions 665–1684):

```
Qy  665  TACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGGAACAA 724
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  605  TACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGGAACAA 664

Qy  725  ATGCTGATGATGGGAATGGTGTGGGAATCAAATAACTAATACAGATCCTATTG 784
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  665  ATGCTGATGATGGGAATGGTGTGGGAATCAAATAACTAATACAGATCCTATTG 724

Qy  785  CTCAAACTAAAACCACTACTGATAATCAAAATCTTCAACTTTTAATTCAGGAGCAATGC 844
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  725  CTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGC 784

Qy  845  CTGGTGCAACAATAGATACGATCCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 904
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  785  CTGGTGCAACAATAGATACGATCCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 844

Qy  905  TCCAATTAGATGAAAAATTGTTATCCAGAATGACTGGTTCTGAAGAGAATAAAAAATA 964
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  845  TCCAATTAGATGAAAAATTGTTATCCAGAATGACTGGTTCTGAAGAGAATAAAAACAGA 904

Qy  965  TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAACAGATATTGGATTCTTGACATAC 1024
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  905  TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAACAGATATTGGATTCTTGACATAC 964

Qy  1025 CCGGGACTCCACAAGTTACTTAAAAGAAGATCAGTTAACGTATTTCAAGACTATACT 1084
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  965  CCGGGACTCCACAAGTTACTTAAAAGAAGATTCAGTTAACGTATTTCAAGACTATACT 1024

Qy  1085 TAAACTCAGTTAATTCTTTATCAATTCATTGGTGATAGTATTTATATTTTTTGGTACCTCTG 1144
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1025 TAAACTCAGTTAATTCTTTATCAATTCATTGGTGATAGTATTTATATTTTTTGGTACCTCTG 1084

Qy  1145 AATTACCATCATTATGGTACTATTCATTCCCAACTAGATTAATCTGATCATGATCAACAAACG 1204
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1085 AATTACCATCATTATGGTACTATTCATTCCCAACTAGATTAATCTGATCATGATCAACCGCTCA 1144

Qy  1205 ATCAAGTTAAAACAGACATCTGATACATCTGAGTTGTTCAAACAGGTGCTGGAACAGGAAATACTA 1264
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1145 ATCAAGTTAAAACAGACATCTGATACATCTGATGTGTTCAAACAGGTGCTGGAACAGGAAATACTA 1204

Qy  1265 GAACAACGACACAACGCTGATCATCTAGTGTTCAAACAGGTGCTGGAACAGGAAATACTA 1324
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1205 GAACAACGACAACAACGCTGATCATCTAGTGTTCAAACAGGTGCTGGAACAGGACAGACTA 1264

Qy  1325 CTAACACTTCTCAAACAGTTTCTAATCCTACTTTCTAAATCGTATCGTATACTTATCGGAATTG 1384
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1265 CTCGACTTCTCAAACAGTTTCTAATCCTACTTTCTAAATCGTATCGTATACTTATCGGAATTG 1324

Qy  1385 ATAGTAAACCAACTTCTGCAAACAAATAGATGAAAACTAATTGGGCACGATCCTAACGGTA 1444
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1325 ATCAAGTAAACCAACTTCTGCAAACAAATAGATGAAAACTAATTGGGCACGATCCTAACGGTA 1384

Qy  1445 TTGAAGCAAGAATATATGCTGAATACAGATTAGCTGTATTCAAAATTGAAATTCCAATAACTA 1504
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1385 TTGAAGCAAGAATATATGCTGAATACAGATTAGCTGTATTCAAAATTGAAATTCCAATAACTA 1444

Qy  1505 ATGCAGGAAACTTTATCCGAAACAACAATTGGTGTGTTGGTTTTACTTCAACAGGTTCAA 1564
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1445 ATGCAGGAAACTTTATCCGAAACAACACAATTGGTGTGTTGGTTTTTACTTCAACAGGTTCAA 1504

Qy  1565 GAGTAGTTTTAAGAGCTTCTATAACGGTGATCAACGTCCAACTGGAAACTTCCAACCTT 1624
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1505 GAGTAGTTTTAAGAGCTTCTATAACGGTGATCAACGTCCAACTGGAAACTTCCAACCTT 1564

Qy  1625 TCTTTATACGTATTTGGTTATTTAGGATACCAACAAACTAGAACAACAAAACACATTCTTTATA 1684
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1565 TCTTTATACGTATTTGGTTATTTAGGATAACCAACACAACTAGAACAACAAAACTTTCTTCGGTACG 1624

Qy  1685 GAACATATAAGCTTTAAAACAACAGCCCTTACGACGTATTAGATTTCTCCAGAGTAGGTA 1744
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1625 GAACATATAAGCTTTACAGAACACAGCCCTTACGACGTATTAGATGCTGCAAGAGTAGGTA 1684
```

CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3159 BP; 1017 A; 639 C; 647 G; 856 T; 0 U; 0 Other;

Query Match 3.9%; Score 124.6; DB 7; Length 3159;
Best Local Similarity 63.5%; Pred. No. 6.9e-17;
Matches 190; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
Qy 2770 TTTGGGGCCTTACCTTCATGGGTAGTGCTTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
        ||||                     |||||||        ||  ||||||||||||||||
Db 2797 TTTGCAGCACTCCCTGCATGGGTGATCCCTGTATCAGTAGTTCTTCAGTTGGGATCTTG 2856

Qy 2830 GCAATTATGATCATCTTAGGATTAGCTACGGTATTCCTTTAAGAGCTCAAAGAAAATTA 2889
        |||||||        |||||||||   ||| ||||||   |||||  |||||||
Db 2857 TTTATCTTGTTAGTCTTAGGACTTGGGATTGGGAGTTAAGAAACTC 2916

Qy 2890 CAAGACAAAGGGTTCAAACAACATTCAAAAAGTGATACCTTGACTGCTGCTGTTGGT 2949
        |||||   |       ||| ||||||||||| |||||||    ||  |||| |||||
Db 2917 CAAGATGCATCCGTTTGTTAATGTCTTTAAAAAGGTTGATACCACTCCAACTGCTGTCGGT 2976

Qy 2950 TCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTGCTGCTTTA 3009
        | |||||||||  |||||||||||||||||  |  |||| ||  |||| ||||| |||
Db 2977 AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGCTGCTAAAAAGCCACCTAGTGCATTG 3036

Qy 3010 GGTGCTGGTAAATCTGGTGATAAGAAGAAACCTGCTCGCTGCTAAACCTGCTCCAGC 3068
        | ||||| ||| |||  ||||||||||||||||||||    |||||||   ||||||
Db 3037 AAAGCTGCTAATCCTAGTGTTAAAAAACCTGCTGCTTTTTTAAAACCACCTGTTCAACC 3095
```

RESULT 7
AAT58840
AAT58840_2
Continuation (3 of 6) of AAT58840 from base 200001 (Mycoplasma genitalium genome.)
WP Sequence split into 6 fragments LOCUS AAT58840 Accession Aat58840

WP	Fragment Name	Begin	End
WP	AAT58840_0	1	110000
WP	AAT58840_1	100001	210000
WP	AAT58840_2	200001	310000
WP	AAT58840_3	300001	410000
WP	AAT58840_4	400001	510000
WP	AAT58840_5	500001	580073

Query Match 3.9%; Score 124.6; DB 2; Length 110000;
Best Local Similarity 63.5%; Pred. No. 2e-16;
Matches 190; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
Qy 2770 TTTGGGGCCTTACCTTCATGGGTAGTGCTTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
        ||||                     |||||||        ||  ||||||||||||||||
```

```
Db 2765 TTCTTGCAATTATGATCATCTTAGGATTAGCTTAGGCTATTCGTAGTGGTATTCCTTTAAGAGAA 2824
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2885 AATTACAAGACAAAGGGTTCAAACACATTCAAAAAGTGATACCTTGATACGTGCTGCTG 2944
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2825 AATTACAAGACAAAGGGTTCAAACACATTCAAAAAGTTGATACCTTGATACGTGCTGCTG 2884
Qy 2945 TTGGTTCAGTTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTGCTG 3004
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2885 TTGGTTCAGTTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTGCTG 2944
Qy 3005 CTTTAAGTGCTGGTAAATCTGGTGATAAGAAGAAACCTGCTCGCTGCTAAACCTGCTCTC 3064
        ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2945 CTTTAAGGTGCTGGTAAATCTGGTGATAAGAAGAAACCTGCTCGCTGCTAAACCTGCTCTC 3004
Qy 3065 CAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAATCTG 3124
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3005 CAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAATCTG 3064
Qy 3125 GTGCGCCTACAAAACCAACTGCTCCTAGCCAGCTGCTCCAAAACCAACCGCTCCCAAAG 3184
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3065 GTGCGCCTACAAAACCAACTGCTCCTAGCCAGCTGCTCCAAAACCAACCGCTCCCAAAG 3124
Qy 3185 AATAA 3189
        |||||
Db 3125 AATAA 3129
```

RESULT 6
ACA39446
ID ACA39446 standard; DNA; 3159 BP.
XX
AC ACA39446;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #21103.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycoplasma genitalium.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU35576.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 27316; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid

```
Db   28703 TTTGCAGCACTCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG 28762
                 |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Qy    2830 GCAATTATGATCATCTTAGGATTAGCTATCCGTATTCCTTTAAGAGAGCTCAAAGAAAATTA 2889
                 ||||||||| | || | |||||||||  ||||| ||  ||||||  |||||  |||
Db   28763 TTTATCTTGTTAGTCTTAGGACTTGGGATTCGGGATCCCAATGTACAGGGTAAGAAAACTC 28822
                 |||||||| || |  |||||||||||||||||||| |||| |
Qy    2890 CAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTTGCTGTTGGT 2949
                 |||| ||||||||||||||||||||||||||||||||||||||||||  |||
Db   28823 CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAACTGCTGTCGGT 28882
                 |||  |   ||  |||||||  |||||||| ||||||
Qy    2950 TCAGTTTACAAGAAGATTATTACCCAAACTGCTACACGTTAAGAAAAAACCTGCTGCTTTA 3009
                 ||| || |||| |||||||| |||||||||| ||| ||||||||  |||||  |
Db   28883 AGTGTGTACAAAAAGATTATTACCCAAACTCGTGTGTGGTGAAAAAAGCACCTAGTGCATTG 28942
                 ||| ||||||||||||||  |||||||  | |  | |||| |  |||| || |||
Qy    3010 GGTGCGTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCTGCTGCTCCAGC 3068
                 |||||||| ||  ||||||||| |||||  ||||||
Db   28943 AAAGCTGCTARATCCTAGTGTGTTAAAAAACCTGCTGCTTTTTTAAAACCACCTGTTCAACC 29001
```

```
RESULT 8
AAQ81778
ID   AAQ81778 standard; DNA; 4182 BP.
XX
AC   AAQ81778;
XX
DT   09-AUG-1995  (first entry)
XX
DE   DNA encoding cytadhesin protein.
XX
KW   Cytadhesin protein; diagnostic probe; unusual codon usage; immunoassay;
KW   antigen; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key            Location/Qualifiers
FT   CDS            713..4081
FT                  /*tag= a
FT                  /product= "cytadhesin protein"
FT   misc_difference 887..889
FT                  /*tag= b
FT                  /codon= seq:tga, aa:trp
FT   misc_difference 998..1000
FT                  /*tag= c
FT                  /codon= seq:tga, aa:trp
FT   misc_difference 1484..1486
FT                  /*tag= d
FT                  /codon= seq:tga, aa:trp
FT   misc_difference 1487..1489
FT                  /*tag= e
FT                  /codon= seq:cta, aa:lys
FT   misc_difference 1559..1561
FT                  /*tag= f
FT                  /codon= seq:tta, aa:lys
FT   misc_difference 1592..1594
FT                  /*tag= g
FT                  /codon= seq:tta, aa:lys
FT   misc_difference 1676..1678
FT                  /*tag= h
FT                  /codon= seq:tta, aa:lys
FT   misc_difference 1715..1717
FT                  /*tag= i
FT                  /codon= seq:tga, aa:trp
FT   misc_difference 1853..1855
FT                  /*tag= j
FT                  /codon= seq:ctt, aa:lys
FT   misc_difference 1890..1892
FT                  /*tag= k
FT                  /codon= seq:tga, aa:trp
FT   misc_difference 1904..1906
FT                  /*tag= l
FT                  /codon= seq:tga, aa:trp
FT   misc_difference 1961..1963
FT                  /*tag= m
FT                  /codon= seq:tta, aa:lys
```

```
FT   misc_difference 2132..2134
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= n
FT   misc_difference 2237..2239
FT                  /codon= seq:ctt, aa:lys
FT                  /*tag= o
FT   misc_difference 2273..2275
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= p
FT   misc_difference 2417..2419
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= q
FT   misc_difference 2504..2506
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= r
FT   misc_difference 2609..2611
FT                  /codon= seq:tta, aa:lys
FT                  /*tag= s
FT   misc_difference 2639..2641
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= t
FT   misc_difference 2858..2860
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= u
FT   misc_difference 3020..3022
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= v
FT   misc_difference 3083..3085
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= w
FT   misc_difference 3138..3140
FT                  /codon= seq:tga, aa:trp
FT                  /*tag= x
XX
PN   US5378820-A.
XX
PD   03-JAN-1995.
XX
PF   09-NOV-1992;  92US-00973257.
XX
PR   09-NOV-1992;  92US-00973257.
XX
PA   (DOHM/) DOHMS J B.
PA   (KEEL/) KEELER C L.
XX
PI   Keeler CL,  Dohms JE;
XX
DR   WPI; 1995-051314/07.
DR   P-PSDB; AAR64927.
XX
PT   acid encoding cytadhesin protein - used as a probe to diagnose Mycoplasma
PT   gallisepticum infection in poultry.
XX
PS   Claim 1; Col 7-16; 13pp; English.
XX
CC   The sequence encodes a cytadhesin protein from Mycoplasma gallisepticum.
CC   All or part of the sequence may be used as a probe for diagnosis of M.
CC   gallisepticum infection in poultry, e.g. turkey or fowl. The sequence has
CC   several UGA codons, which may be eliminated to avoid problems of
CC   expression in vitro due to irregular codon usage. The sequence may be
CC   cloned in a vector for recombinant cytoadhesin production, for use as an
CC   antigen in immunoassay formats
XX
SQ   Sequence 4182 BP; 1456 A; 763 C; 746 G; 1217 T; 0 U; 0 Other;

Query Match     2.5%;  Score 78.4;  DB 2;  Length 4182;
Best Local Similarity   98.8%;  Pred. No.7.5e-07;
Matches   79;  Conservative   0;  Mismatches   1;  Indels   1;  Gaps   0;

Qy      1 ATGAATATTTCTAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4103 ATGAATATTTCTAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTTGGA 4162
```

```
Qy    61 GCTCTTGGTTCTCGCAAGCTT 80
            ||||||||||||||||||||
Db  4163 ACTCTTGGTTCTCGCAAGCTT 4182
```

RESULT 9
AAQ79746
ID   AAQ79746 standard; DNA; 3435 BP.
XX
AC   AAQ79746;
XX
DT   25-MAR-2003 (revised)
DT   08-SEP-1995 (first entry)
XX
DE   Mycoplasma pirum adhesin gene.
XX
KW   Adhesin; cytoadherence; Mycoplasma; human immunodeficiency virus; HIV;
KW   vaccine; infection; ss.
XX
OS   Mycoplasma pirum.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..3435
FT   /tag= a
FT   /codon= TGA; aa:Trp
FT   /note= "typical of Mycoplasma"
XX
PN   FR2705970-A1.
XX
PD   09-DEC-1994.
XX
PF   04-JUN-1993;   93FR-00006745.
XX
PR   04-JUN-1993;   93FR-00006745.
XX
PA   (INSP) INST PASTEUR.
XX
PI   Tham TN, Ferris S, Blanchard A, Montagnier L, Bahraoui E;
XX
DR   WPI; 1995-024735/04.
DR   P-PSDB; AAR76059.
XX
PT   Nucleic acid encoding Mycoplasma pirum adhesin - and related proteins and
PT   antibodies, useful in diagnosis, treatment and prevention of M.pirum and
PT   HIV infection.
XX
PS   Claim 2; Fig 1; 55pp; French.
XX
CC   The M.pirum adhesin gene was isolated using a probe based on the 3'- end
CC   of the known M.genitalium adhesin gene. M.pirum is found in AIDS patients
CC   and the adhesin may have a role in infection of cells by HIV. If so, the
CC   M.pirum adhesin protein will be useful in vaccines to protect against the
CC   cytopathic effect of the Mycoplasma and against that of HIV. (Updated on
CC   25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 3435 BP; 1334 A; 466 C; 494 G; 1141 T; 0 U; 0 Other;

Query Match   2.0%; Score 64.6; DB 2; Length 3435;
Best Local Similarity   48.3%; Pred. No. 0.00068;
Matches 181; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

Qy 2727 TAAATTAAAGAACAATAAATTAGTAACATTCCTTATAACAACTTTCGCGCCTTACCTTC 2786
            ||| ||| ||| || |||| |||||| | ||| |||||||| | ||
Db 3033 TAGTTCTAAAATAAATGTTATAAAAATTCTATTTTGCTGGATTAGTGCAATGGCTGA 3092

Qy 2787 ATGGGTAGTGCCTACAGCAAGAATTGGTAGTACATTAGGTATTCTTGCAATTATGATCATCTT 2846
            | |||| | | ||||| ||||| | | ||||||| || |||| || |||
Db 3093 TTGAATTCTTCCTGTAGTTATTGCTATTCCCATTGTATTAGTTGCATTAATAATTGGTTT 3152

Qy 2847 AGGATTAGCTATCGGTATTCCTTAAGAGCTCAAAGAAAATTACAAGACAAAGGGTTCAA 2906
            |||| ||| |||||| || ||||| ||| | ||||||
Db 3153 AGGTTGCAGTATTGGGATTCCAATGGCTAAACATAAAAAACATAAAAAGCTATTAAAGT... 3212
```

```
Qy 2907 AACAACATTCAAAAAGTTGATACCTTGACGCGCTGCTGCTGGTTGGTTCAGTTACAAGAAGAT 2966
            |||||| || |||  |||||||| |||  |  ||| | ||||| ||||||||  |||||||
Db 3213 ATTGCAACACGATAAAGTTGAACATTGACTTCAGCTGTGTGGTGTTTCAAAAAAT 3272

Qy 2967 TATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTCGTCTTTAGGTGCTGGTAAATCTGG 3026
            |||||||  ||| |  ||||||||||  |  ||
Db 3273 TATTGACAATACAAATTCTAATAATGTGAAATCTAAACCACAAATGTTAGGCGTGCAGC 3332

Qy 3027 TGATAAGAAACCTGCTGCTGCTCGTCGTCGTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAA 3086
            |||||  |  ||||| | |||||||||||  |  |||||||||| |||||||||||||
Db 3333 TAAAAAACCAAATACAGTTCCACCTGCTAGATCTCAATTAACAAATGATTCTGTTTCTAG 3392

Qy 3087 AGCTAGTCACCAGC 3101
            ||||||||||||
Db 3393 ACCCACTCCACCATC 3407
```

RESULT 10
ABQ75107/c
ID   ABQ75107 standard; cDNA; 4985 BP.
XX
AC   ABQ75107;
XX
DT   01-NOV-2002 (first entry)
XX
DE   Anopheles gambiae gambiae odourant receptor 2 genomic DNA SEQ ID NO:10.
XX
KW   Anopheles gambiae; mosquito; olfactory gene; arrestin 1; pest control;
KW   odourant receptor; olfaction; gene; ds.
XX
OS   Anopheles gambiae.
XX
PN   WO200259274-A2.
XX
PD   01-AUG-2002.
XX
PF   28-JAN-2002; 2002WO-US002549.
XX
PR   26-JAN-2001; 2001US-0264649P.
PR   24-JAN-2002; 2002US-00056405.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Zwiebel LJ;
XX
DR   WPI; 2002-627421/67.
DR   P-PSDB; ABP52835.
XX
PT   New mosquito olfaction polypeptides and polynucleotides, useful for
PT   mosquito management, i.e. controlling the pest and disease vectors, or
PT   for identifying pest control agents.
XX
PS   Disclosure; Fig 4a; 96pp; English.
XX
CC   The present invention describes a purified Anopheles gambiae olfaction
CC   polypeptide comprising a 383, 394, 380, 411, 412, 391, 157 or 401 residue
CC   amino acid sequence (see ABP52833 to ABP52840) (S1), a conservatively
CC   modified amino acid sequence of them, or a sequence of (S1) with at least
CC   20 consecutive residues. Also described: (1) an isolated polynucleotide
CC   comprising: (a) a nucleotide sequence encoding the purified Anopheles
CC   gambiae olfaction polypeptide; or (b) a nucleotide sequence that
CC   hybridises under stringent conditions to a hybridisation probe comprising
CC   a 1964, 1239, 1142, 1236, 1194, 1176, 474 or 1206 nucleotide sequence
CC   (see ABQ75102 to ABQ75105 and ABQ75110 to ABQ75113) (S2), or its
CC   complement; and (2) a method for identifying an agent that binds to
CC   mosquito olfaction molecule comprising: (a) providing an isolated
CC   mosquito olfaction molecule; (b) contacting a test agent with the
CC   isolated mosquito olfaction molecule; and (c) detecting specific binding
CC   of the test agent to the isolated mosquito olfaction molecule, where the
CC   presence of specific binding identifies the test agent as a mosquito
CC   olfaction-binding compound. The mosquito olfaction molecules are useful
CC   for mosquito management, i.e. controlling this pest and disease vector. A
```

```
CC   method from the present invention of screening for substances that
CC   modulate arrestin-odourant receptor interaction is useful for identifying
CC   pest control agents. The present sequence represents Anopheles gambiae
CC   odourant receptor 2 genomic DNA from the present invention. N.B. The
CC   features given in figure 4a are tentative and do not directly encode SEQ
CC   ID NO:6 (ABP52835)
XX
SQ   Sequence 4985 BP; 1319 A; 1058 C; 990 G; 1618 T; 0 U; 0 Other;

Query Match      2.0%;     Score 62.6; DB 6;  Length 4985;
Best Local Similarity 45.5%;  Pred. No. 0.0021;
Matches 333;  Conservative    0;  Mismatches 396;  Indels    3;  Gaps    3;

Qy     253 ATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGTGAGTGATTGTGGT 312
           |||  ||||| |  || |   |||| | ||| |||  ||| || ||||| ||| || ||||
Db    3070 ATAAAAAAAACAATAATAACATAATAAAGAATAACAACAACAACAATATATAAGAA 3011

Qy     313 GGATACAAAGTAAAAACAAATGTTCA-GATTACCAACTAGCAGAAATAGATTTGATCA 371
           |   || | |  ||||  || ||||  |||| |||| ||   |  ||||||  |||
Db    3010 CAACAACAACAATAATAATATAATAACAATAATAACAATAATAAGAATAATAA 2951

Qy     372 AAGACAAACAAGAGCATATTGCTCTGTTGGTTAATGATGAAGCTAACGTTCAT-TTAA 430
           |  || || ||| ||||| |||  | |   ||||||| ||  |||||  ||| |||
Db    2950 TAACAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 2891

Qy     431 AAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTGTAA 490
           |||| |||||| ||   ||| || ||| || ||  |||| || |||||  ||| ||
Db    2890 AAAAAATAATTATTATATAAGAATAGTAATAATAATTAAAAATAATAATATTA 2831

Qy     491 TTGGTGGTGTTGATAATCCACGTAATTACTTACTGATGATGGGACTAAATTTA 550
           ||||| || ||  |||| || ||||| |||||||| | |||| || | ||||
Db    2830 ATAATAATAATAATAATAATAATAAGAATAATAATAATAATAATAATAATAATA 2771

Qy     551 ATTTTACAAACCAAACTCAAGGTGAAATTGTTAATAATGTTAATGACTTCATTTTAGATGCGCCAATCT 610
           ||||  || || || |||||||| ||| ||  ||   ||| |||| | ||||| ||  | ||  ||||
Db    2770 ATAATAATAATAATCATAATAAAATTATTATTATTATGATAATAGTAATAATGAAT 2711

Qy     611 TAC-CTAAAGATTTACACCCAGATTGGTATAACTTATAACATTCAAAGAAGAGATCTTACCA 669
           ||| |||| |  |||| ||||| |  || |||  || |||  ||| || |  || | ||
Db    2710 AACAATAATAATTAATTAATTATAAAATAATAATAATAATATATAATATAATATAT 2651

Qy     670 AATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTGAGTAGTGGAACAAATGCT 729
           || |||| ||||| ||| || || |||||| ||||| ||||| ||||||| |||
Db    2650 TATTATTGTTTATAATAATAATAATAATAATAATCATAATAATTCATAATAT 2591

Qy     730 GATGATGGGGATGTTTGATTGTGGGAAGTGTCAAATAACTAATACAGATCCTATTGCTCAA 789
           |||||||||| ||| ||| | ||| ||||| | ||| |||| |||| ||  ||||| |||
Db    2590 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 2531

Qy     790 ACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGCCTGGT 849
           ||| | |  ||| || |||||  |  | | |||| ||||| || ||   |  ||  |||
Db    2530 AGTAATAATTAATTATTAGTAATAATAATAATAATAATATTAATAATAAT 2471

Qy     850 GCAAACAATAGATACGATTCTCAAATTGAAATGTCAAGCATAGAATAAAACATCTTCCAA 909
           ||  |    || || ||||  ||| || ||| ||||| || | ||| | | ||| |||
Db    2470 AATAATAATAATAATAATAATATTATTATTATAATAATAATAAT 2411

Qy     910 TTAGATGAAAAATTTGTTTATCCAGAATGACTGGTTCTGAAGAATAAAAATATTACA 969
           ||||||| |   ||||| ||   |  | ||| |||| | |||||| |  | ||  ||
Db    2410 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 2351

Qy     970 AGATTAGCTACT 981
           |||||   | | |
Db    2350 AAATTCTGTATT 2339

RESULT 11
ACF79720/c
ID   ACF79720 standard; DNA; 4985 BP.
XX
AC   ACF79720;
XX
DT   15-JAN-2004  (first entry)
```

```
XX
DE   Mosquito odorant receptor 2 genomic DNA.
XX
KW   Mosquito; odorant receptor 2; olfaction; insecticide; antimalarial; gene;
KW   ss.
XX
OS   Anopheles gambiae.
XX
FH   Key             Location/Qualifiers
FT   exon            1..702
FT                   /*tag= d
FT                   /number= 1
FT   5'UTR           1..485
FT                   /*tag= a
FT   TATA_signal     295..300
FT                   /*tag= b
FT   CDS             486..3561
FT                   /*tag= c
FT                   /product= "Odorant receptor 2"
FT                   /note= "contains introns"
FT   intron          703..798
FT                   /*tag= e
FT                   /number= 1
FT   exon            799..882
FT                   /*tag= f
FT                   /number= 2
FT   intron          883..978
FT                   /*tag= g
FT                   /number= 2
FT   exon            979..1404
FT                   /*tag= h
FT                   /number= 3
FT   intron          1405..1601
FT                   /*tag= i
FT                   /number= 3
FT   exon            1602..1702
FT                   /*tag= j
FT                   /number= 4
FT   intron          1703..1783
FT                   /*tag= k
FT                   /number= 4
FT   exon            1784..1886
FT                   /*tag= l
FT                   /number= 5
FT   intron          1887..1993
FT                   /*tag= m
FT                   /number= 5
FT   exon            1994..2106
FT                   /*tag= n
FT                   /number= 6
FT   intron          2107..3459
FT                   /*tag= o
FT                   /number= 6
FT   exon            3460..4985
FT                   /*tag= p
FT   3'UTR           3562..4985
FT                   /*tag= q
XX
PN   WO200307659O-A2.
XX
PD   18-SEP-2003.
XX
PF   10-MAR-2003; 2003WO-US007174.
XX
PR   08-MAR-2002; 2002US-00094240.
XX
PA   (UIVA-) UNIV VANDERBILT.
XX
PI   Zwiebel LJ;
XX
DR   WPI; 2003-722331/68.
XX
PT   New mosquito arrestin 1 and 2 genes and polypeptides, useful for
```

PT identifying mosquito olfaction molecule binding compounds which reduce
PT the ability of mosquitoes to locate sources of bloodmeal, e.g. humans.
XX
PS Disclosure; Fig 4a; 101pp; English.
XX
CC The present sequence is that of genomic DNA encoding a novel mosquito
CC olfaction molecule, odorant receptor 2. A cDNA sequence for odorant
CC receptor 2 is given in ACF79719. The invention provides 9 novel mosquito
CC polypeptides and the nucleic acids encoding them. These are odorant
CC receptor molecules 1-7 and arrestins 1 and 2. The odorant receptors
CC function in a ligand-induced signal transduction pathway for the
CC activation of mosquito olfaction. Arrestin functions to inhibit the
CC activated signal transduction cascade. Thus, the odorant receptors act as
CC an 'on' switch, and arrestin as an 'off' switch for the odorant detection
CC system of the mosquito. Methods are provided for identifying compounds
CC that interfere with the operation of the mosquito olfactory system,
CC particularly compounds that modulate arrestin 2 activity. These are
CC useful for the control of mosquitoes, particularly by reducing their
CC ability to locate sources of bloodmeal
XX
SQ Sequence 4985 BP; 1319 A; 1058 C; 990 G; 1618 T; 0 U; 0 Other;

Query Match   2.0%; Score 62.6; DB 9; Length 4985;
Best Local Similarity   45.5%; Pred. No. 0.0021;
Matches 333; Conservative   0; Mismatches 396; Indels   3; Gaps   3;

Qy   253 ATCAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTAAGTCATGATGATTGTGGT 312
          ||||||| |||||| ||||||||||||| |||||||||||||||||| |||||| |||||
Db   3070 ATAAAAAAAAACATTAATAATAAACAATATAACAATATAACACAACAACAATAATAAGAA 3011

Qy   313 GGATACAAAGTAAAACAAATAGTTTCA-GATTACACAACTAGCACAGAAATAGATTTGATCA 371
          ||||||| |||| |||||| ||||| ||||||| |||||||||| |||||| |||||
Db   3010 CAACACACAACAACAATAATAAGAATATAACAATATAACAATAGCAATTAATAATAAGAATAATAA 2951

Qy   372 AAGACAAACAAGAGCATATTATGTCTCGTTGGTTAATGATGAGCTAACGTTCAT-TAA 430
          |||||||| |||||| |||||| |||||||| |||||| |||||||||||| | |||
Db   2950 TAACAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 2891

Qy   431 AAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTGTAA 490
          |||||||||| | |||||||||||| |||||||||||| ||||||||||||||||||
Db   2890 AAAAAAATAATTATTATATAATAATAAGAATAGTAATAATAATTATTAAAAAATAATATTA 2831

Qy   491 TTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAAATTTA 550
          |||||||||||||||| ||||| ||||| ||||||||| |||||| |||||||| |||
Db   2830 ATAATAATAATATATAATAATAAGAATAATAATAATATATAATAATAATAATAATAATA 2771

Qy   551 ATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCT 610
          |||||||||||||| |||||||||||| ||||| |||||||||||| |||||| |||||
Db   2770 ATAATAATAATAATCATAATAAATTATTATTATGATAATAATAATGTATAATAATGAAT 2711

Qy   611 TAC-CTAAAGATTTACACCCAGTTGGTATAACTTATCAAAGAAAGATCTTACCA 669
          ||| |||||||||||||||||||| ||||||||| ||||||||||||| |||||
Db   2710 AACATAATAATAATTAATTAAATAAATTAATTAATAATAATTAATAATAATAAT 2651

Qy   670 AATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTGGAACAAAATGCT 729
          |||||||| ||||| ||| ||||||||||| |||||||||| ||||||||||||
Db   2650 TATTATTGTTATATAATAATAATAATAATAATCATAATTCATAATTATAAT 2591

Qy   730 GATGATGGGGATTGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTAGTGCTCAA 789
          ||||||||| |||||||||| |||||| ||| |||||||| ||||||| |||||||| ||||
Db   2590 AATAATAATAAAATAATAATAGTAATAATAATAATATATAATAATAATAAATAATAAA 2531

Qy   790 ACTAAAACCACTACTGATAATCAAAATCCTCAACTTTTAATTCAGGAGCAATGCCTCGGT 849
          |||||||||||| |||||||| ||||| || ||||| ||||||| |||||||| ||||
Db   2530 AGTAATAATAATAATAATAGTAATAATAATAATATATAATAATAATAATAATAATAAT 2471

Qy   850 GCAAACAATAGATACGACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTTCCAA 909
          ||||||||| |||| |||||||| |||||||||||| |||||||||||| ||||||||||
Db   2470 AATAATAATAATAATAATTAATTAATTATTAATAATAATAATAATAATAATAATAATAAT 2411

Qy   910 TTAGATGAAAAATTTGTTTATCCAGAATGGACTCGTCGTTGGTTAATGAAGAGAATATTACA 969
          |||||||||| |||||| ||||||| ||||| ||||| ||||||| ||||||||||||||
Db   2410 AATAATAATAATAATAATAATAATAATAATATATAATAATAATAATAATAATAATAAT 2351

Qy   970 AGATTAGCTACT 981
          ||||||||||||
Db   2350 AAATTCTGTATT 2339

RESULT 12
ABL33825/c
ID   ABL33825 standard; DNA; 8771 BP.
XX
AC   ABL33825;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1798.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2002009288-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 1798; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 8771 BP; 2734 A; 59 C; 1660 G; 4318 T; 0 U; 0 Other;

Query Match   1.9%; Score 60.8; DB 6; Length 8771;
Best Local Similarity   46.3%; Pred. No. 0.006;
Matches 200; Conservative   0; Mismatches 232; Indels   0; Gaps   0;

Qy   256 AAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGTGATGATGATGTGGTGGA 315
          |||||||||||||| |||||||||| ||||| ||||||| ||||||||||| ||||||
Db   5053 ATAATCTATACATACAATAATAATATTATTCAACTTTAAAAAAAAAAAAAAAAACTTAATCTA 4994

Qy   316 TACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCACAACTAGCAGAAATAGATTTGATCAAGA 375
          |||||||| |||| ||||||||||||| |||||| ||||| ||||||| |||||
Db   4993 TACTACCACAAAAAAAAAAACCTTAAAAACATTATTAAAAAACATTATCAAAATAAAATAAACCAATAACAAAAA 4934

Qy   376 CAAACAAGAGCATATTATGTCTCGTTGGTTAATGATGGTTGGTTAATGAAGCTAACGTTCATTTAATGAAGA 435
          ||||

Db 4933 AATACACTATATAAATTCCACTTCTAAAAACTACCTAAAAATAATCAAAATTATAAAAAC 4874

Qy 436 ATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTTGTAATTGGT 495

Db 4873 AAAAATAATAACATAAATTCAAAAACCTAAAAAAAAAAAATAATAAAAAATTATTTATTAAT 4814

Qy 496 GGTGTTGATAATCCAGCTCACGTAATTAGATTTACTCGATGATGGGACTAAATTTAATTTT 555

Db 4813 AATAATAATAATATACAACAAAAATATAATAAACATAATTAATTACCACTAAATTATCACA 4754

Qy 556 ACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGATGCGCCAATCTTACCT 615

Db 4753 TACAAATAATCAAATAATAATAAATTTTATATTATCATTATATATTTACCA 4694

Qy 616 AAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCTTACCAAATGAC 675

Db 4693 TAATACAAAAAATAATAAAAAAAAAAAAAAAAAAAAATAAAACTTAACGATTACATAACTTA 4634

Qy 676 GTCAACACTGCA 687

Db 4633 ACCAATACCTCA 4622

RESULT 13
AAA30290/c
ID AAA30290 standard; DNA; 3489 BP.
XX
AC AAA30290;
XX
DT 15-SEP-2003 (revised)
DT 11-SEP-2000 (first entry)
XX
DE Kaposi's sarcoma-associated herpesvirus LANA gene.
XX
KW Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
KW latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
KW Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;
KW Kaposi's sarcoma; primary effusion lymphoma; PEL;
KW human immunodeficiency virus; HIV; multicentric Castleman's disease; ds.
XX
OS Human herpesvirus 8.
XX
FH Key   Location/Qualifiers
FT CDS   1..3489
FT   /*tag= a
FT   /product= "LANA"
FT misc_signal   40..50
FT   /*tag= b
FT   /note= "nuclear localisation signal, NLS"
FT misc_signal   190..210
FT   /*tag= c
FT   /note= "nuclear localisation signal, NLS"
XX
PN WO200029626-A1.
XX
PD 25-MAY-2000.
XX
PF 19-NOV-1999; 99WO-US027508.
XX
PR 19-NOV-1998; 98US-00109422.
PR 21-APR-1999; 99US-00298568.
XX
PA (KIEF/) KIEFF E D.
PA (BALL/) BALLESTAS M E.
PA (KAYE/) KAYE K M.
XX
PI Kieff ED, Ballestas ME, Kaye KM;
XX
DR WPI; 2000-387829/33.
DR P-PSDB; AAY96255.
XX
PT Treating or preventing a disease associated with rhodino virus infection
PT in a mammal which includes Kaposi's Sarcoma and Primary Effusion

PT Lymphoma.
XX
PS Disclosure; Fig 6; 70pp; English.
XX
CC The present sequence is the Kaposi's sarcoma-associated herpesvirus,
CC (KSHV) latency-associated nuclear antigen (LANA) gene. KSHV is also known
CC as Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or gamma
CC -2 herpes virus class. The LANA protein is necessary for the efficient
CC persistence of rhadino virus DNA in mammalian cells. Persistent rhadino
CC virus infection is implicated in a variety of diseases e.g. Kaposi's
CC Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
CC Castleman's disease. In addition, KS is a common malignancy in HIV
CC patients. KSHV persists in host cells in a latent form. One of the few
CC genes expressed from the latent viral DNA is LANA. LANA associates with
CC both human chromosomes and with the rhadino virus cis-acting element
CC (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
CC "tied" to the host chromosomes. This allows the viral DNA to persist in
CC the host cell. The present sequence may be used to screen and identify
CC molecules that inhibit LANA interaction with RVCAE, thereby interfering
CC with the latency cycle of this virus. Potential antiviral treatments for
CC the above mentioned diseases may therefore be based on LANA deregulation.
CC (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;

Query Match   1.9%;   Score 60.2;   DB 3;   Length 3489;
Best Local Similarity   61.6%;   Pred. No. 0.0062;
Matches 114;   Conservative 0;   Mismatches 68;   Indels 3;   Gaps 1;

Qy 2998 CCTGCTGCTTAGGTGCTGTGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT 3057
Db 2236 CATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTCATCCT 2177

Qy 3058 GCTGCTCCAGCTAAACCATCTGACCACAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
Db 2176 GCTGCTGCTCGTCGTCATCCTGCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTCGTCCT 2117

Qy 3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCTGCTGCTCCAAAACCAACC 3174
Db 2116 GCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCATCCTGCT 2057

Qy 3175 GCTCC 3179
Db 2056 GCTGC 2052

RESULT 14
AAF82901/c
ID AAF82901 standard; DNA; 3489 BP.
XX
AC AAF82901;
XX
DT 11-SEP-2003 (revised)
DT 06-AUG-2003 (revised)
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of KSHV tethering protein, LANA.
XX
KW Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KW Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW KSHV; latency-associated nuclear antigen; LANA; ds.
XX
OS Human herpesvirus 8.
XX
FH Key   Location/Qualifiers
FT CDS   1..3489
FT   /*tag= a
XX
PN WO200125484-A2.
XX
PD 12-APR-2001.
XX
PF 29-SEP-2000; 2000WO-US026908.

```
XX
PR    01-OCT-1999;   99US-00410399.
XX
PA    (UNMI ) UNIV MICHIGAN.
XX
PI    Robertson ES, Cotter MA;
XX
DR    WPI; 2001-281736/29.
DR    P-PSDB; AAB62331.
XX
PT    A composition for use in gene therapy comprises an expression vector that
PT    includes a nucleic acid sequence encoding a nucleic acid binding protein.
XX
PS    Disclosure; Fig 9A; 60pp; English.
XX
CC    The invention provides a composition comprising nucleic acid, histone H1
CC    protein and expression vector operationally encoding a protein suitable
CC    for tethering the nucleic acid to the histone H1 protein, where the
CC    tethering protein is LANA. The composition is useful in aiding the
CC    retention of the viral DNA in the host cell. The viral vector encodes a
CC    protein suitable for tethering DNA to Histone H1. Methods for screening
CC    for compounds which are agonistic or antagonistic for the tethering of
CC    viral proteins to histone H1 and DNA binding sites are useful for
CC    developing the method of viral transfer. The composition has applications
CC    to gene therapy, including the treatment of multiple sclerosis,
CC    Parkinson's disease, Huntington disease and diabetes. The present
CC    sequence represents the nucleotide sequence of the Kaposi's sarcoma
CC    associated herpesvirus (human herpesvirus 8) latency-associated nuclear
CC    antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-
CC    2003 to correct OS field.) (Updated on 11-SEP-2003 to standardise OS
CC    field)
XX
SQ    Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;

Query Match          1.9%;   Score 60.2;  DB 4;  Length 3489;
Best Local Similarity   61.6%;   Pred. No. 0.0062;
Matches  114;  Conservative   0;  Mismatches   68;  Indels    3;  Gaps    1;

Qy   2998  CCTGCTGCTTAGGTGCTGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
           |||||||||||||||  |||||||| |||| |||||||||||||||||||||
Db   2236  CATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCATCCT 2177

Qy   3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGC-- 3115
           |||||||| || ||  |||  |||      |||||||||| ||  |||||||||
Db   2176  GCTGCTGCTGCTCATCCTGCTCATCCTGCTCCTGCTGCTGCTCATCCTGCTGCTCCT 2117

Qy   3116  -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGGCTCCAAAAACCAACC 3174
            ||||  || || || ||||||||||||||||| |  |||||||||||  ||||||
Db   2116  GCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTCATCCTGCTCATCCTGCTCATCCTGCT 2057

Qy   3175  GCTCC 3179
           |||||
Db   2056  GCTGC 2052
```

RESULT 15
```
ABA93487/c
ID    ABA93487 standard; DNA; 3489 BP.
XX
AC    ABA93487;
XX
DT    25-APR-2002  (first entry)
XX
DE    Kaposi's sarcoma-associated herpesvirus LANA protein encoding DNA.
XX
KW    Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAB; PEL;
KW    KSHV terminal repeat; rhadino virus cis acting element; episome;
KW    primary effusion lymphoma; latency-associated nuclear antigen;
KW    gene therapy; gene transfer; gene; ds.
XX
OS    Human herpesvirus 8.
XX
FH    Key           Location/Qualifiers
```

```
XX
FT    CDS              1..3489
FT                     /*tag= a
FT                     /product= "LANA protein"
FT                     /note= "latency-associated nuclear antigen"
XX
PN    US6322792-B1.
XX
PD    27-NOV-2001.
XX
PF    21-APR-1999;   99US-00298558.
XX
PR    19-NOV-1998;   98US-0109422P.
XX
PA    (KIBF/) KIEFF E D.
XX
PI    Kieff ED,  Ballestas ME,  Kaye KM;
XX
DR    WPI; 2002-153769/20.
DR    P-PSDB; ABB05621.
XX
PT    System for episomal retention of plasmids in mammalian cells, useful in
PT    gene therapy, comprises rhadinoviral LANA and RVCAB sequences.
XX
PS    Claim 1; Fig 6; 27pp; English.
XX
CC    The present invention describes a system (A) for maintaining a plasmid as
CC    an episome in mammalian cells, comprising the rhadinoviral sequence LANA
CC    (latency-associated nuclear antigen) of 3489 base pairs (see ABA93487,
CC    S1) expressed in the cell, and the rhadinoviral sequence RVCAB
CC    (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2)
CC    present in the plasmid. Also describes a method for maintaining a
CC    closed circular DNA in a cell by expressing (S1) in the cells and having
CC    (S2) as a cis-acting and maintenance sequence in the DNA. (A) is
CC    particularly used in gene therapy (or other gene transfer applications)
CC    that uses mammalian cells in which LANA is expressed. (A) improves
CC    persistence of gene therapy vectors in cells. The present sequence
CC    encodes Kaposi's sarcoma-associated herpesvirus (KSHV, also called human
CC    herpesvirus 8) LANA protein, which is used in the exemplification of the
CC    present invention
XX
SQ    Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;

Query Match          1.9%;   Score 60.2;  DB 6;  Length 3489;
Best Local Similarity   61.6%;   Pred. No. 0.0062;
Matches  114;  Conservative   0;  Mismatches   68;  Indels    3;  Gaps    1;

Qy   2998  CCTGCTGCTTAGGTGCTGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
           |||||||||||||||  |||||||| |||| |||||||||||||||||||||
Db   2236  CATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCATCCT 2177

Qy   3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGC-- 3115
           |||||||| || ||  |||  |||      |||||||||| ||  |||||||||
Db   2176  GCTGCTGCTGCTCATCCTGCTCATCCTGCTCCTGCTGCTGCTCATCCTGCTGCTCCT 2117

Qy   3116  -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGGCTCCAAAAACCAACC 3174
            ||||  || || || ||||||||||||||||| |  |||||||||||  ||||||
Db   2116  GCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTCATCCTGCTCATCCTGCTCATCCTGCT 2057

Qy   3175  GCTCC 3179
           |||||
Db   2056  GCTGC 2052
```

Search completed: May 5, 2004, 12:25:53
Job time : 1125.68 secs

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May  5, 2004, 10:52:31 ; Search time 240.505 Seconds
                                    (without alignments)
                                    7358.420 Million cell updates/sec

Title:          US-09-901-572A-2
Perfect score:  3189
Sequence:       1 atgaatattctaaaaact........caaccgctcccaagataa 3189

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued Patents NA:*
        1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
        2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
        3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
        4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
        5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
        6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Query Length	DB	ID	Description
	1	124.6	3.9	580073	4	US-08-545-528D-1	Sequence 1, Appli
	2	78.4	2.5	4182	4	US-07-973-257-1	Sequence 1, Appli
	3	60.2	1.9	3489	2	US-08-728-323A-1	Sequence 1, Appli
c	4	60.2	1.9	3489	4	US-09-298-568-1	Sequence 1, Appli
c	5	60.2	1.9	3489	4	US-09-410-399-1	Sequence 1, Appl
	6	60.2	1.9	32207	2	US-08-770-379-20	Sequence 20, Appl
	7	60.2	1.9	32207	4	US-08-757-669A-20	Sequence 20, Appl
	8	60.2	1.9	32207	3	US-09-230-371A-20	Sequence 20, Appl
	9	58.4	1.8	640681	4	US-09-790-988-1	Sequence 1, Appli
c	10	54.2	1.7	832	4	US-09-621-976-2813	Sequence 2813, Ap
	11	53.2	1.7	987	4	US-09-142-584-5	Sequence 5, Appli
	12	53.2	1.7	1037	4	US-09-181-585-3	Sequence 3, Appli
	13	53.2	1.7	1159	4	US-09-181-585-1	Sequence 1, Appli
	14	53.2	1.7	1471	4	US-09-181-585-2	Sequence 2, Appli
	15	53	1.7	987	4	US-09-142-584-3	Sequence 3, Appli
	16	52.4	1.6	3057	4	US-09-601-198-55	Sequence 55, Appli
	17	51.4	1.6	987	4	US-09-142-584-1	Sequence 1, Appli
c	18	51	1.6	636	3	US-08-998-416-1137	Sequence 1137, Ap
	19	51	1.6	1575	4	US-09-543-681A-615	Sequence 615, App
c	20	49.8	1.6	300	4	US-09-328-352-4065	Sequence 4065, Ap
c	21	49.4	1.5	1887	4	US-09-601-198-39	Sequence 39, Appl
	22	49	1.5	5163	3	US-08-700-651-1	Sequence 1, Appli
	23	49	1.5	5163	3	US-08-928-361B-4	Sequence 4, Appli
	24	49	1.5	5163	3	US-09-588-995A-4	Sequence 4, Appli
	25	49	1.5	5318	3	US-08-700-651-2	Sequence 2, Appli
	26	49	1.5	5318	3	US-08-928-361B-3	Sequence 3, Appli
	27	49	1.5	5318	4	US-09-588-995A-3	Sequence 3, Appli
c	28	48.4	1.5	397	3	US-09-253-691-3	Sequence 3, Appli
	29	48.4	1.5	4884	1	US-07-665-792B-10	Sequence 10, Appli
	30	48.2	1.5	234	1	US-08-469-802B-3	Sequence 3, Appli
	31	48.2	1.5	234	1	US-08-267-803B-3	Sequence 3, Appli
	32	47.6	1.5	195	1	US-08-469-802B-2	Sequence 2, Appli
	33	47.6	1.5	195	1	US-08-267-803B-2	Sequence 2, Appli
c	34	47.4	1.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	35	47.2	1.5	165	3	US-09-043-303-17	Sequence 17, Appl
c	36	47.2	1.5	203	3	US-09-043-303-7	Sequence 7, Appl
c	37	47.2	1.5	429	3	US-09-543-681A-707	Sequence 707, App
	38	46.2	1.4	5511	3	US-08-928-361B-2	Sequence 2, Appli
	39	46.2	1.4	5511	4	US-09-588-995A-2	Sequence 2, Appli
	40	46.2	1.4	7334	3	US-08-928-361B-1	Sequence 1, Appli
	41	46.2	1.4	7334	4	US-09-588-995A-1	Sequence 1, Appli
c	42	46	1.4	1588	4	US-09-490-291-7	Sequence 7, Appli
	43	46	1.4	5181	4	US-08-257-073-10	Sequence 10, Appli
	44	46	1.4	16442	3	US-08-781-891-208	Sequence 208, App
	45	46	1.4	16442	4	US-09-618-166-208	Sequence 208, App

ALIGNMENTS

```
RESULT 1
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmen
; Patent No. 6537773
; TITLE OF INVENTION:  Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

  Query Match             3.9%;  Score 124.6;  DB 4;  Length 580073;
  Best Local Similarity   63.5%;  Pred. No. 4e-19;
  Matches  190;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy       2770 TTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
                                            ||||||                        |||
Db     228703 TTTGCAGCACTCCCCTGCATGGGTGATCCCTGTATCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG 228762

Qy       2830 GCAATTATGATCATCTTAGGATTAGCTATCGGTATCTTTAAGAGCTCAAAGAAAATTA 2889
                                            ||||||                       |||
Db     228763 TTTATCTTGTTAGTTCTTAGGACTTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC 228822

Qy       2890 CAAGACAAAGGGTTCAAACAACAACATTCAAAAAAGTTGATACCTTGACTGCGCGTGTTGGT 2949
                  |||                       ||||||
Db     228823 CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAACTGTCGTGTCGGT 228882

Qy       2950 TCAGTTTACAAGAGAAGATTATTACCCAAACTGCTAACGTTAAGAAACAAACTGCTGCTTTA 3009
                                            ||||||
Db     228883 AGTGTGTACAAAAAGATTATTACCCAAACTGCTGTGTGGTGAAAAAAGCACCTAGTGCATTG 228942

Qy       3010 GGTGCTGGTAAATCTGGTGATAAGAAACCTGCTCGTCGCTAAACCTGCTAAACCGCTCCAGC 3068
                                            ||||||
Db     228943 AAAGCTGCTAATCCTAGTGTTAAAAAACCTGCTGCTTTTTAAAACCACCTGTTCAACC 229001

RESULT 2
US-08-545-528D-1
```

US-07-973-257-1
; Sequence 1, Application US/07973257
; Patent No. 5378820
; GENERAL INFORMATION:
; APPLICANT: Keeler, Jr. Calvin L.
; APPLICANT: Dohms, John E.
; TITLE OF INVENTION: Gene Encoding Cytadhesin
; TITLE OF INVENTION: Protein of Mycoplasma Gallisepticum and Its Use
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Connolly and Hutz
;   STREET: 1220 Market Street
;   CITY: Wilmington
;   STATE: Delaware
;   COUNTRY: U.S.A.
;   ZIP: 19899
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM/PC or Compatible
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/973,257
;   FILING DATE: 19921109
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION DATA: No. 5378820e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 4182 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: double stranded
;   TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
;   ORGANISM: Mycoplasma Gallisepticum
;   STRAIN: S6
US-07-973-257-1

Query Match   2.5%; Score 78.4; DB 1; Length 4182;
Best Local Similarity 98.8%; Pred. No. 5e-09;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4103 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 4162

Qy   61 GCTCTTGGTTCTGCAAGCTT   80
         ||||||||||||||||||||
Db 4163 ACTCTTGGTTCTGCAAGCTT 4182

RESULT 3
US-08-728-323A-1/c
; Sequence 1, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Cooper & Dunham LLP
;   STREET: 1185 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.

ZIP: 10036
COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/728,323A
 FILING DATE:
 CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-278-0400
 TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 3489 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
  NAME/KEY: CDS
  LOCATION: 1..3489
US-08-728-323A-1

Query Match   1.9%; Score 60.2; DB 2; Length 3489;
Best Local Similarity 61.6%; Pred. No. 0.0001;
Matches 114; Conservative 0; Mismatches 68; Indels 3; Gaps 1;

Qy 2998 CCTGCTGCTTAGCTGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
         ||||||| |||| |||||||||||||||||||||||||||||||||||||||||
Db 2236 CATCCTGCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTGCTCATCCT 2177

Qy 3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
         |||||||||||||||||||||||||||||||||||||||||||||| ||
Db 2176 GCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCCT 2117

Qy 3116 -CTAAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCACC 3174
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 2116 GCTCATCCTGCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCT 2057

Qy 3175 GCTCC 3179
         ||||
Db 2056 GCTGC 2052

RESULT 4
US-09-298-568-1/c
; Sequence 1, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballestas, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; CURRENT FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
　ADDRESSEE: Cooper & Dunham LLP
　STREET: 1185 Avenue of the Americas
　CITY: New York
　STATE: New York
　COUNTRY: U.S.A.
　ZIP: 10036
COMPUTER READABLE FORM:
　MEDIUM TYPE: Floppy disk.
　COMPUTER: IBM PC compatible
　OPERATING SYSTEM: PC-DOS/MS-DOS
　SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
　APPLICATION NUMBER: US/08/770,379
　FILING DATE:
　CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
　NAME: White, John P.
　REGISTRATION NUMBER: 28,678
　REFERENCE/DOCKET NUMBER: 52342
TELECOMMUNICATION INFORMATION:
　TELEPHONE: (212) 278-0400
　TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 20:
　SEQUENCE CHARACTERISTICS:
　LENGTH: 32207 base pairs
　TYPE: nucleic acid
　STRANDEDNESS: double
　TOPOLOGY: linear
　MOLECULE TYPE: DNA (genomic)
US-08-770-379-20

```
Query Match         1.9%;  Score 60.2;  DB 2;  Length 32207;
Best Local Similarity  61.6%;  Pred. No. 0.00026;
Matches  114;  Conservative  0;  Mismatches  68;  Indels  3;  Gaps  1;

Qy   2998 CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
          |||||||||||||| |  ||||  |||||||||||||   |||||||||||||||||||
Db  19761 CATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTCACCAGCTAAACCAACTGGGCTCATCCT 19820

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
          ||||||  |||||||  | ||||   ||||||||| |||||||||||||||| |||||||
Db  19821 GCTGCTGCTGCTCTCATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCCCT 19880

Qy   3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
          ||||| ||||| |||  ||||||||||||||||||| |||| | |||||||||||||||||
Db  19881 GCTCATCCTGCTGCTCATCCTGCTCGTGCTGCTCATCCTGCTGCTCATCCTGCT 19940

Qy   3175 GCTCC 3179
          |||||
Db  19941 GCTGC 19945
```

RESULT 7
US-08-757-669A-20
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INFORMATION:
;　APPLICANT: Chang, Yuan
;　APPLICANT: Bohenzky, Roy A.
;　APPLICANT: Russo, James J.
;　APPLICANT: Edelman, Isidore S.
;　APPLICANT: Moore, Patrick S.
;　TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
;　TITLE OF INVENTION: SEQUENCES AND USES THEREOF
;　NUMBER OF SEQUENCES: 20
;　CORRESPONDENCE ADDRESS:
;　ADDRESSEE: Cooper & Dunham LLP
;　STREET: 1185 Avenue of the Americas
;　CITY: New York
;　STATE: New York

```
Query Match         1.9%;  Score 60.2;  DB 4;  Length 3489;
Best Local Similarity  61.6%;  Pred. No. 0.0001;
Matches  114;  Conservative  0;  Mismatches  68;  Indels  3;  Gaps  1;

Qy  2998 CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
         |||||||||||||| |  ||||  |||||||||||||   |||||||||||||||||||
Db 2236 CATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTCACCAGCTAAACCAACTGGGCTCATCCT 2177

Qy  3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
         ||||||  |||||||  | ||||   ||||||||| |||||||||||||||| |||||||
Db 2176 GCTGCTGCTGCTCTCATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCCCT 2117

Qy  3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
         ||||| ||||| |||  ||||||||||||||||||| |||| | |||||||||||||||||
Db 2116 GCTCATCCTGCTGCTCATCCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCT 2057

Qy  3175 GCTCC 3179
         |||||
Db 2056 GCTGC 2052
```

RESULT 5
US-09-410-399-1/c
; Sequence 1, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
;　APPLICANT: Robertson, Erle S.
;　APPLICANT: Cotter, Murray A.
;　TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
;　TITLE OF INVENTION: to Genomic Host DNA
;　FILE REFERENCE: UM-03778
;　CURRENT APPLICATION NUMBER: US/09/410,399
;　CURRENT FILING DATE: 1999-10-01
;　NUMBER OF SEQ ID NOS: 6
;　SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;　LENGTH: 3489;
;　TYPE: DNA
;　ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

```
Query Match         1.9%;  Score 60.2;  DB 4;  Length 3489;
Best Local Similarity  61.6%;  Pred. No. 0.0001;
Matches  114;  Conservative  0;  Mismatches  68;  Indels  3;  Gaps  1;

Qy  2998 CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
         |||||||||||||| |  ||||  |||||||||||||   |||||||||||||||||||
Db 2236 CATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTCACCAGCTAAACCAACTGGGCTCATCCT 2177

Qy  3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
         ||||||  |||||||  | ||||   ||||||||| |||||||||||||||| |||||||
Db 2176 GCTGCTGCTGCTCTCATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCCCT 2117

Qy  3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
         ||||| ||||| |||  ||||||||||||||||||| |||| | |||||||||||||||||
Db 2116 GCTCATCCTGCTGCTCATCCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCT 2057

Qy  3175 GCTCC 3179
         |||||
Db 2056 GCTGC 2052
```

RESULT 6
US-08-770-379-20
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
;　APPLICANT: Chang, Yuan
;　APPLICANT: Bohenzky, Roy A.
;　APPLICANT: Russo, James J.
;　APPLICANT: Edelman, Isidore S.
;　APPLICANT: Moore, Patrick S.
;　TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED

COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,669A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 32207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-669A-20

Query Match 1.9%; Score 60.2; DB 3; Length 32207;
Best Local Similarity 61.6%; Pred. No. 0.00026; Gaps 1;
Matches 114; Conservative 0; Mismatches 68; Indels 3;

Qy 2998 CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTCTAAACCT 3057
Db 19761 CATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCATCCT 19820

Qy 3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
Db 19821 GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCCTGCTGCTCCT 19880

Qy 3116 -CTAAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
Db 19881 GCTCATCCTGCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCT 19940

Qy 3175 GCTCC 3179
Db 19941 GCTGC 19945

US-09-230-371A-20

RESULT 8
US-09-230-371A-20
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A
APPLICANT: Russo, James J
APPLICANT: Edelman, Isidore S
APPLICANT: Moore, Patrick S
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 45185-G-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,371A
CURRENT FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: PCT/US97/13346
PRIOR FILING DATE: 1997-07-22
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match 1.9%; Score 60.2; DB 4; Length 32207;
Best Local Similarity 61.6%; Pred. No. 0.00026; Gaps 1;
Matches 114; Conservative 0; Mismatches 68; Indels 3;

Qy 2998 CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTCTAAACCT 3057
Db 19761 CATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCATCCT 19820

Qy 3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
Db 19821 GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCCTGCTGCTCCT 19880

Qy 3116 -CTAAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
Db 19881 GCTCATCCTGCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCT 19940

Qy 3175 GCTCC 3179
Db 19941 GCTGC 19945

RESULT 9
US-09-790-988-1
; Sequence 1, Application US/0979098 8
; Patent No. 6632935
; GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEMI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 1.8%; Score 58.4; DB 4; Length 640681;
Best Local Similarity 47.3%; Pred. No. 0.0026; Gaps 1;
Matches 209; Conservative 0; Mismatches 231; Indels 2;

Qy 241 AATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATGGTTTAGAATCTAAGT 300
Db 525047 ATTAAAAAAAAATAAACACAGAAAATTTAATACATACCATTCTAAAAGTAATTTTATT 525106

Qy 301 GATGATTGTGGTGGATACAAAGTAAAACAAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
Db 525107 TTAAAAAAGAAGAAGAAAATATTATAAAACTATTATCTCAAAAAAGAATATAAAAAAA 525166

Qy 361 AGATTTGATCAAAGACAAACAAGAGACATATATGCTCTGTTGGTTAATGATGAAGCTAAC 420
Db 525167 GCAATTGTTAAAATAGATCCTTCAATCTATAAAAAAATATAACAAATCAAGAAGCACAA 525226

Qy 421 GTTCATTTAAAAGAAGAATAAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT 480
Db 525227 ATTTATTTCAAAAAAAATCAAGATAATTTTTATTTTAAAGTAATTAATTTT 525286

Qy 481 AAGTTTGTAATTGGTGGTTGATAATCCAGCTCACGTAACTAGATTTACTGATGATGGG 540
Db 525287 GTTGAATTAAAAACTGATAACTTTAAAATACATTGCGAAAATAAAGAGATTTATGATTGG 525346

Qy 541 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTTCATTTTAGAT 600
Db 525347 TATATAAGAAAATATTAC--ACAATACTCAACTAAAGAAAAAGAAGATATAGTATTATTC 525404

Qy 601 GCGCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG 660

Db   525405 AAGTCAAAAATAAAACAACAGCCATATCGATATTATCTAGATTGCATAATACACCAGAAG 525464
                    ||| |||| ||||| |||||| | || || || ||||| |||
Qy   661 ATCTTACCAAATGACGTCAACA 682
                 || || || |||| |||||
Db   525465 ATTTTTCAAAAATAGCTCAAGA 525486

RESULT 10
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match   1.7%; Score 54.2; DB 4; Length 832;
Best Local Similarity 13.1%; Pred. No. 0.0015; Indels   0; Gaps   0;
Matches   47; Conservative 162; Mismatches 150; Indels   0; Gaps   0;

Qy   1348 AATCCTACTTTAAATACTTATCGTAGTTTTGGAATTGATAGTAAACCAACTTCTGCAAAC 1407
              ||| ||||| |||| : ::| :::: ||| || :: ||| : | :
Db   390 AATTGTACCATAAAATATATTTTTGTYWTWNKTWNYWTTYTWTRMWWMKKKARRWYY 331

Db   1408 AAAATAGAATGAAACTAATTGGGCAGATCCTAACGTTATTGAAGCAAGAATATATGCTGAA 1467
              :: | ||| : : :::: :|:||: : || : | |:|: |
Qy   330 WWKSTYACASRYRKYTWGWWYMYWKRMMSTRWYCYMCWKCCMVRGRRCAWYTWARGRMWS 271

Qy   1468 TACAGATTAGGTATTCAAAATGAAATTCCAATAACTAATGCAGGAAACTTTATCCGAAAC 1527
              : :: | || | : : |: | | :: | : : :::|: | :|:
Db   270 YAWGKWKSMRSAMSMCTRMYKKGSTYWTMKCTCATWCYWYWKYWKRMWSKTCWSGSRGG 211

Qy   1528 ACAATTGTGCTGTCTGTTGGTTTTTACTTCAAGAGTTCAAGAGAGTAGTTTTAAGAGCTTCTTAT 1587
              |: ::||||| ||| | :| | |::: :::: ::| ::| || |:
Db   210 YMTSYSYTSRSYSMYVWASWMYTMCWWWGRWWSTYWTMAWGKKWWRYATTWRRAMMWWWAAW 151

Qy   1588 AACGGTGATCAACGTCCAACTGGAAACTTCCAACCTTCTTCTTAATACGTATTTGGTTATTTA 1647
              || | ::: || || | :: :|: : :::|| :: : |:|:| :: | :|
Db   150 TMMWYMWWAWCMSSRGAAMYRRTMMMGYRYWWRKKSYRRTRCAWYAYAWKTKRSYYWCWR 91

Qy   1648 GGATACCAACAAACTAGAACAGGAACTTCTCGGTACGGAACATATAAGCTTTTAAACAA 1706
              ::|: |::::::: ::|:| || ::| |: :::: :|:: |:
Db   90 WKWKRCMWMMMAMAYGKTMMRACWKTRYWRWWAWAMWRMWTMWMMYYWYWRAMKRR 32

RESULT 11
US-09-142-584-5
; Sequence 5, Application US/09142584
; Patent No. 6403094
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Havard, Helen L.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Payne, Dean W.
; TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
; FILE REFERENCE: 124-665
; CURRENT APPLICATION NUMBER: US/09/142,584

; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/GB97/00660
; EARLIER FILING DATE: 1997-03-11
; EARLIER APPLICATION NUMBER: GB 9605222.0
; EARLIER FILING DATE: 1996-03-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Clostridium perfringens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (136)..(987)
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: (1)..(32)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(984)
; FEATURE:
; OTHER INFORMATION: "n" at positions 451-453 represent a, t, c, g or
; OTHER INFORMATION: other
US-09-142-584-5

Query Match   1.7%; Score 53.2; DB 4; Length 987;
Best Local Similarity 44.3%; Pred. No. 0.0028; Indels   0; Gaps   0;
Matches   208; Conservative   0; Mismatches 261; Indels   0; Gaps   0;

Qy   363 ATTTGATCAAGACAAACAAGAGCATATTATGCTCGTTGGTTAATGATGAGCTAACGT 422
              || ||| || ||||| || ||| || : : ||:| || || || |||||||
Db   117 ATCTAATGAAATGTCCAAAAAAGCTTCTTATGATAATGTAGATACATTGAGAGAAGG 176

Db   423 TCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAA 482
              || || |||||||| ||| |||||||||||| ||: | |||| || | ||||||
Qy   177 AAGATATAATACAAATAATAATTACTTAAAGAATGAAAAATATTATCCTAATGCTAT 236

Qy   483 GTTGTAATTGGTGTGTTGATAATCCAGTCACGTAATCCAAGAGGAA 542
              | ||||| || || ||||| | || || || |: ||| |||||
Db   237 GGCATATTTTGATAAGGTTACTATAAAATCCACAAGGAAATGATTTTTATATTAATAATCC 296

Db   543 TAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGC 602
              ||| | |: |::||| || | ||||| ||||| |:|| |: || :| |||:
Qy   297 TAAAGTTGAATTAGATGGAGGAACCATCAATGAATTATCTTGAAGATGTTTATGTTGGAAA 356

Qy   603 GCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGAT 662
              ||| || :||||| || ||:| || |||| ||||| ||||| |: | |||||| ||
Db   357 AGCTCTTCTTAACTAATGATACTCAACAAGAACAAAAATTAAAATCACAATCATTCACTTG 416

Db   663 CTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTTAGTGGAAC 722
              || | || | ::| ||| | :|| | | :| :: || | | |:|| |||||
Qy   417 TAAAAATACTGATACAGTACCTGCAACTACTACTNNNACTGTGGGAACTTCGATACAAGC 476

Qy   723 AAATGCTGATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTAT 782
              ||||| |||| | |:|| ::| ||| :|:|| ||| |||| || :|| |||||
Db   477 AACTGCTAAGTTTACTGTTCCTTTTAATGAAACAGGAGTATCATCACTACTAGTTATAG 536

Qy   783 TGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTAAT 831
              ||| ||| ||| | |||| |:|||||| || | | | :: |||
Db   537 TTTTGCAAATACAAATACAAATACTAATTCAAAAGAAATTACTTCATAT 585

RESULT 12
US-09-181-585-3
; Sequence 3, Application US/09181585
; Patent No. 6524791
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 1100900101
; CURRENT APPLICATION NUMBER: US/09/181,585
; CURRENT FILING DATE: 1998-10-28

```
;   NUMBER OF SEQ ID NOS: 18
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 3
;   LENGTH: 1037
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: cDNA
;   OTHER INFORMATION: comprising exons E, C, and A
US-09-181-585-3

Query Match          1.7%;    Score 53.2;  DB 4;  Length 1037;
Best Local Similarity  50.8%;  Pred. No. 0.0028;
Matches  127;  Conservative   0;  Mismatches 123;  Indels   0;  Gaps   0;

Qy   2930 CCTTGACTGCTGCTGTTGGTTCAGTTTACAAGAAGAGATTATTACCCAAACTGCTAACGTTA 2989
               ||||||            |||||       ||||        ||||
Db    605 CCTGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA  664

Qy   2990 AGAAAAAACCTGCTACTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGCTG 3049
                |            |||||||||||||||||||||
Db    665 CTACTACTACTGCTACTGCTGCTGCTGCTGCTGCTG  724

Qy   3050 CTAAACCTGCTGCTCCAGCTAAACCATCTGACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
               ||||            |||||||||||
Db    725 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  784

Qy   3110 CTGGGGCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTGCTGCTG 3169
               |||||||||
Db    785 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  844

Qy   3170 CAACCGCTCC 3179
               |||||||
Db    845 CTGCTGCTGC  854

RESULT 13
US-09-181-585-1
;   Sequence 1, Application US/09181585
;   Patent No. 6524791
;   GENERAL INFORMATION:
;   APPLICANT: Ranum, Laura P.W.
;   APPLICANT: Koob, Michael
;   TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
;   FILE REFERENCE: 11000900101
;   CURRENT APPLICATION NUMBER: US/09/181,585
;   CURRENT FILING DATE: 1998-10-28
;   NUMBER OF SEQ ID NOS: 18
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1
;   LENGTH: 1159
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-181-585-1

Query Match          1.7%;    Score 53.2;  DB 4;  Length 1159;
Best Local Similarity  50.8%;  Pred. No. 0.003;
Matches  127;  Conservative   0;  Mismatches 123;  Indels   0;  Gaps   0;

Qy   2930 CCTTGACTGCTGCTGTTGGTTCAGTTTACAAGAAGAGATTATTACCCAAACTGCTAACGTTA 2989
               ||||||            |||||        ||||        ||||
Db    417 CCTGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA  476

Qy   2990 AGAAAAAACCTGCTACTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGCTG 3049
                |            ||||||||||||||
Db    477 CTACTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTG  536

Qy   3050 CTAAACCTGCTGCTCCAGCTAAACCATCTGACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
               ||||            |||||||||
Db    537 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  596

Qy   3110 CTGGGGCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTGCTGCTG 3169
               ||          ||||||
```

```
Db    597 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  656

Qy   3170 CAACCGCTCC 3179
               |||||||
Db    657 CTGCTGCTGC  666

RESULT 14
US-09-181-585-2
;   Sequence 2, Application US/09181585
;   Patent No. 6524791
;   GENERAL INFORMATION:
;   APPLICANT: Ranum, Laura P.W.
;   APPLICANT: Koob, Michael
;   TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
;   FILE REFERENCE: 11000900101
;   CURRENT APPLICATION NUMBER: US/09/181,585
;   CURRENT FILING DATE: 1998-10-28
;   NUMBER OF SEQ ID NOS: 18
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 2
;   LENGTH: 1471
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: cDNA
;   OTHER INFORMATION: comprising exons D, C, B, and A
US-09-181-585-2

Query Match          1.7%;    Score 53.2;  DB 4;  Length 1471;
Best Local Similarity  50.8%;  Pred. No. 0.0033;
Matches  127;  Conservative   0;  Mismatches 123;  Indels   0;  Gaps   0;

Qy   2930 CCTTGACTGCTGCTGTTGGTTCAGTTTACAAGAAGAGATTATTACCCAAACTGCTAACGTTA 2989
               ||||||            |||||        ||||        ||||
Db   1039 CCTGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA 1098

Qy   2990 AGAAAAAACCTGCTACTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGCTG 3049
                |            |||||||||||||||
Db   1099 CTACTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTG 1158

Qy   3050 CTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
               ||||            ||||||||
Db   1159 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218

Qy   3110 CTGGGGCTAAATCTGGTGCGGCCTACACAAAACCAACTGCTCCTAAGCCAGCTGCTGCTCCAAAAC 3169
               |||||||                 |||||
Db   1219 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278

Qy   3170 CAACCGCTCC 3179
               |||||||
Db   1279 CTGCTGCTGC 1288

RESULT 15
US-09-142-584-3
;   Sequence 3, Application US/09142584
;   Patent No. 6403094
;   GENERAL INFORMATION:
;   APPLICANT: Titball, Richard W.
;   APPLICANT: Williamson, Ethel D.
;   APPLICANT: Havard, Helen L.
;   APPLICANT: Oyston, Petra C.F.
;   APPLICANT: Payne, Dean W.
;   TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
;   FILE REFERENCE: 124-665
;   CURRENT APPLICATION NUMBER: US/09/142,584
;   CURRENT FILING DATE: 1998-09-11
;   EARLIER APPLICATION NUMBER: PCT/GB97/00660
;   EARLIER FILING DATE: 1997-03-11
;   EARLIER APPLICATION NUMBER: GB 9605222.0
;   EARLIER FILING DATE: 1996-03-12
;   NUMBER OF SEQ ID NOS: 6
```

```
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 3
;  LENGTH: 987
;  TYPE: DNA
;  ORGANISM: Clostridium perfringens
;  FEATURE:
;  NAME/KEY: mat_peptide
;  LOCATION: (136)..(987)
;  FEATURE:
;  NAME/KEY: misc_signal
;  LOCATION: (1)..(32)
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(984)
US-09-142-584-3
```

```
Query Match             1.7%;    Score 53;  DB 4;  Length 987;
Best Local Similarity  44.6%;    Pred. No. 0.0031;
Matches  209;  Conservative   0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy     363 ATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGT 422
               |||||||  ||||||  |||||  |||       ||  ||||  ||||||
Db     117 ATCTAATGAAATGTCCAAAAAAGCTTCTTATGATAAATGTAGATACATTAATTGAGARAAGG 176

Qy     423 TCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAA 482
                |   |||||||  ||||  | ||| |  | | |||||||  |||| ||||||
Db     177 AAGATATAATACAAAATATTAATTACTTAAAGAGAATGGAAAAATATTATCCTAATGCTAT 236

Qy     483 GTTTGTAATTGGTGGTGTTGTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGAC 542
               ||||           ||||||||  ||||||  ||||||||||||||
Db     237 GGCATATTTTGATAAGGTTACTATAAATCCACAAGGAAATGATTTTTATATTAATAATCC 296

Qy     543 TAAATTTAATTTTACAAACCAAACTCAAGGTGAAAATTGTTAATGACTTCATTTTAGATGC 602
               ||||||  || ||||||  ||| ||||||||||||  || ||   |||||||
Db     297 TAAAGTTGAATTAGATGGGAGAACCATCAATGAATTATCTTGAGAGATGTTTATGTTGGAAA 356

Qy     603 GCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGAT 662
                | |||||||  ||||||  |||   |   ||| ||| ||||  ||||| ||||
Db     357 AGCTCTTCTTAACTAATGATACTCAACAAGAACAAAAATTAAAATCACAATCATTCACTTG 416

Qy     663 CTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTTAGTGGAAC 722
               ||||||||||  || |||||||   ||||  ||||  || |||||||  ||
Db     417 TAAAAATACTGATACAGTAGCTGCCAACTACTACTCCGACTGTGGGAACTTCGATACAAGC 476

Qy     723 AAATGCTGATGATGGGATGTTTGATTGTGGGATGGTCAAATAACTAATACAGATCCTAT 782
               |||||||||||  ||||   ||| |||| |||||  |  ||||||||||  ||||
Db     477 AACTGCTAAGTTTACTGTGTTCCTTTTAATGAAACAGGAGTATCATTAACTACTAGTTATAG 536

Qy     783 TGCTCAAACTAAAAACCACTACTGATAATCAAAATCCTTCAACTTTAAT 831
               |||||||||   |||      | |||||||||   ||||   |||||
Db     537 TTTTGCAAATACAAATACTAATTCAAATACTAATTCAAAAGAAATTACTCATAAT 585
```

Search completed: May 5, 2004, 20:04:15
Job time : 247.505 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May  5, 2004, 17:07:04 ; Search time 1255.74 Seconds
                                      (without alignments)
                                      11489.234 Million cell updates/sec

Title:       US-09-901-572A-2
Perfect score: 3189
Sequence:    1 atgaatatttctaaaaaact........caaccgctcccaaagaataa  3189

Scoring table: IDENTITY NUC
             Gapop 10.0, Gapext 1.0

Searched:    2937390 seqs, 226206796 residues

Total number of hits satisfying chosen parameters:     5874780

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
             Maximum Match 100%
             Listing first 45 summaries
```

Published Applications NA:*
```
 1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 3:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 6:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 7:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

SUMMARIES

```
Result           %     Query
No.    Score   Match  Length  DB  ID                      Description
----   -----   -----  ------  --  --                      -----------
  1    3189   100.0    3189   10  US-09-901-572A-2        Sequence 2, Appli
  2    3189   100.0    3189   15  US-10-131-591A-2        Sequence 2, Appli
  3   3143.4   98.6    8354   14  US-10-125-818-1         Sequence 144, App
  4   3057.8   95.9    3129   15  US-10-131-591A-79       Sequence 79, Appl
  5    124.6    3.9    3159   13  US-10-282-122A-27316     Sequence 27316, A
  6    124.6    3.9  580073   15  US-10-205-220-1         Sequence 1, Appli
c 7     62.6    2.0    4985   15  US-10-094-240-10        Sequence 10, Appl
c 8     62.6    1.9    4985   15  US-10-056-405-10        Sequence 1798, Ap
c 9     60.8    1.9    8771   13  US-10-311-455-1798       Sequence 1, Appli
c 10    60.2    1.9    3489   13  US-09-894-273-1         Sequence 1, Appli
c 11    60.2    1.9    3489    9  US-09-294-804-1         Sequence 1, Appli
  12    58.4    1.8  640681    9  US-09-790-988-1         Sequence 1, Appli
c 13    53.6    1.7 3673778   15  US-10-312-841-1          Sequence 3, Appli
c 14    53.2    1.7    1037   16  US-10-373-667-3         Sequence
```

```
     15   53.2  1.7   1159  16  US-10-373-667-1        Sequence 1, Appli
     16   53.2  1.7   1471  16  US-10-373-667-2        Sequence 2, Appli
     17     53  1.7  13784  17  US-10-257-166-144      Sequence 144, App
c    18   52.4  1.6   3057  15  US-10-349-680-148      Sequence 148, App
     19   51.6  1.6    906  13  US-10-282-122A-16807   Sequence 16807, A
     20   51.6  1.6   2451  15  US-10-091-007-55       Sequence 55, Appl
     21   51.4  1.6   2553  13  US-10-282-122A-40676   Sequence 40676, A
     22   50.8  1.6   2017  15  US-10-155-533-3        Sequence 3, Appli
     23   50.8  1.6   5314  15  US-10-155-533-1        Sequence 1, Appli
     24   50.2  1.6   2010   9  US-09-878-756-3        Sequence 3, Appli
c    25   49.8  1.6    486  13  US-10-465-217-13       Sequence 13, Appl
c    26   49.4  1.5   1887  15  US-10-349-680-133      Sequence 133, App
c    27     49  1.6    718  16  US-10-264-049-422      Sequence 422, App
     28   48.6  1.5    357  13  US-10-085-783A-7843    Sequence 7843, Ap
     29   48.6  1.5    357  16  US-10-242-535A-7843    Sequence 7843, Ap
     30     48  1.5    720  13  US-10-424-599-129163   Sequence 129163,
c    31     48  1.5   2183  16  US-10-104-047-1064     Sequence 1064, Ap
     32   47.8  1.5   3931  15  US-10-006-780-1        Sequence 1, Appli
     33   47.6  1.5   1220  15  US-10-169-710-46       Sequence 46, Appl
     34   47.4  1.5    972  13  US-10-243-552-209      Sequence 209, App
     35   47.2  1.5    769  13  US-10-363-616-47       Sequence 47, Appl
     36     47  1.5    303  13  US-10-085-783A-33947   Sequence 33947, A
     37     47  1.5    303  16  US-10-242-535A-33947   Sequence 33947, A
     38     47  1.5    364  13  US-10-085-783A-15800   Sequence 15800, A
     39     47  1.5    364  16  US-10-242-535A-15800   Sequence 15800, A
     40     47  1.5    375  10  US-09-918-995-6127     Sequence 6127, Ap
     41     47  1.5    712   9  US-09-822-849A-186     Sequence 186, App
     42     47  1.5    722   9  US-09-923-304-1        Sequence 1, Appli
     43     47  1.5    722   9  US-09-880-107-1764     Sequence 1764, Ap
c    44     47  1.5    755  13  US-10-027-632-128109   Sequence 128109,
c    45     47  1.6    755  16  US-10-027-632-128109   Sequence 128109,
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

```
RESULT 1
US-09-901-572A-2
; Sequence 2, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: mgc3 gene
US-09-901-572A-2

Query Match             100.0%;  Score 3189;  DB 10;  Length 3189;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  3189;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 ATGAATATTTCTAAAAAACTTAAAAGTATACATTGATAGGTGATTAGCTGTATTGA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGAATATTTCTAAAAAACTTAAAAGTATACATTGATAGGTGATTAGCTGTATTGA 60

Qy        61 GCTCTCGGTTCTGCGCAAGCTTTGGCTTAAGCAATCAGATAGAGTAACGATAACACGCAA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GCTCTCGGTTCTGCGCAAGCTTTGGCTTAAGCAATCAGATAGAGTAACGATAACACGCAA 120

Qy       121 TTAGTTAATCAAGCAAGACGCTAGATGCTAATTCGTTAGACTTGCGCAGGTCTTGGACAA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TTAGTTAATCAAGCAAGACGCTAGATGCTAATTCGTTAGACTTGCGCAGGTCTTGGACAA 180

Qy       181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGAATAACTTTATAACAGCAGCT 240
```

```
Db  1261 AACCGAACACGACAACGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320

Qy  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTACTTAAATACTTATCGTAGTTTTGGA 1380
Db  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTACTTAAATACTTATCGTAGTTTTGGA 1380

Qy  1381 ATTGATAGTAAACCAACTTCTGCAACAACAAATAGATGAAACTAATTGGGCGATCCTAAC 1440
Db  1381 ATTGATAGTAAACCAACTTCTGCAACAACAAATAGATGAAACTAATTGGGCAGATCCTAAC 1440

Qy  1441 GTTATTGAAGCAAGAATATATGCTGATACAGATTAGGTATTCAAAATGAAATTCCAATA 1500
Db  1441 GTTATTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA 1500

Qy  1501 ACTAATGCAGGAAACTTATCCGAAACACAATGGTGTTGTTTTACTTCAACAGGT 1560
Db  1501 ACTAATGCAGGAAACTTATCCGAAACACAATGGTGTTGTGTTTTACTTCAACAGGT 1560

Qy  1561 TCAAGAGTAGTTTTAAGAGCTTCTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA 1620
Db  1561 TCAAGAGTAGTTTTAAGAGCTTCTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA 1620

Qy  1621 CCTTTCTTATACGTATTTGGTATTGGTGATTTAATTACAACAAACTAGAACAGGAACTTCTGG 1680
Db  1621 CCTTTCTTATACGTATTTGGTATTGGTGATTTAATTACAACAAACTAGAACAGGAACTTTCTGG 1680

Qy  1681 TACGGAACATATAGCTTTTAAACAACAGCCCTTACGACGTATTAGATTCTCAAGAGTA 1740
Db  1681 TACGGAACATATAGCTTTTAAACAACAGCCCTTACGACGTATTAGATTCTCAAGAGTA 1740

Qy  1741 GGTACTGAAACCAATCAATTTAGAAGAACTTCATTAACTACTCCATATATAAGAGCACAAGGTGAC 1800
Db  1741 GGTACTGAAACCAATCAATTTAGAAGAACTTCATTAACTACTCCATATATAAGAGCACAAGGTGAT 1800

Qy  1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCCATATATAAGAGCACAAGAGGTGAC 1860
Db  1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCCATATATAAGAGCACAAGAGGTGAC 1860

Qy  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT 1920
Db  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT 1920

Qy  1921 CAATCAGTTTTGATTGAATTGAGATTAGAAATTAGAAAACTTAGAATGTTGGGGTTAAAGCATCA 1980
Db  1921 CAATCAGTTTTGATTGAATTGAGATTAGAAATTAGAAAACTTAGAATGTTGGGGTTAAAGCATCA 1980

Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGTCTAGAAATGATTGCTGCAACAACA 2040
Db  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGTCTAGAAATGATTGCTGCAACAACA 2040

Qy  2041 TACTTAAGATCACAAATTGGATTAGCTGATTCGATTACCAACAACAACCATTC 2100
Db  2041 TACTTAAGATCACAAATTGGATTAGCTGATTCGATTACCAACAACAACCATTC 2100

Qy  2101 GGAACAACTCACCAAGTTATTTCAGTATCAACCTGGTGATCAGTTCAGTTCTCATCAATTAAGAAT 2160
Db  2101 GGAACAACTCACCAAGTTATTTCAGTATCAACCTGGTGATCAGTTCAGTTCTCATCAATTAAGAAT 2160

Qy  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTAGTGTACTTCTTATTCACAAATGAAAATAAT 2220
Db  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTAGTGTACTTCTTATTCACAAATGAAAATAT 2220

Qy  2221 AAATTCAGTTGTTTATACATTAACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280
Db  2221 AAATTCAGTTGTTTATACATTAACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280

Qy  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAAATTGGTGTAATCTTACCTTTATTAGAC 2340
Db  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAAATTGGTGTAATCTTACCTTTATTAGAC 2340

Qy  2341 AATTCATTCTATACAGTAAAATGCTTGCTGGTAAATGCTTGCTGGTAAATGTTCATTGTTCTCATCAAACCCTGT 2400
Db  2341 AATTCATTCTATACAGTAAAATGCTTGCTGGTAAATGCTTGCTGGTAAATGTTCATTGTTCTCATCAAACCCTGT 2400
```

```
Db   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240

Qy   241 AATGGACAATTATCAAATTATAGTAGTTTTACTAAAACCATTATAATGGTTTAGATCTAAGT 300
Db   241 AATGGAACAATTATCAAATTATAGTAGTTTTACTAAAACCATTATAATGGTTTAGATCTAAGT 300

Qy   301 GATGATTGTGGTGGATACAAAGTAAACAAATAGTTTCAGATTACCACAACTAGCAGAGAAAT 360
Db   301 GATGATTGTGGTGGATACAAAGTAAACAAATAGTTTCAGATTACCACAACTAGCAGAGAAAT 360

Qy   361 AGATTTGATCAAGACAAACAAGAGCATATTATGTCTCTGTTGGTTAATGATGAGCTAAC 420
Db   361 AGATTTGATCAAGACAAACAAGAGCATATTATGTCTCTGTTGGTTAATGATGAGCTAAC 420

Qy   421 GTTCATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCT 480
Db   421 GTTCATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCT 480

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGG 540
Db   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGG 540

Qy   541 ACTAAATTTAATTTTACAAACCAAGGTGAAATTGTAATGACTTCATTTAGAT 600
Db   541 ACTAAATTTAATTTTACAAACCAAGGTGAAATTGTAATGACTTCATTTAGAT 600

Qy   601 GCGCCAATCTAAAGATTTACACCCAGATTGGTATAAACTTATACATTCAAAGAAAG 660
Db   601 GCGCCAATCTAAAGATTTACACCCAGATTGGTATAAACTTATACATTCAAAGAAAG 660

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTGGCCAGTAGGTAGAGTTAGTGGA 720
Db   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTGGCCAGTAGGTAGAGTTAGTGGA 720

Qy   721 ACAAATGCTGATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT 780
Db   721 ACAAATGCTGATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT 780

Qy   781 ATTGCTCAAACTAAAACCACTACTGACTAAATCAAATAATCCTTCAACTTTTAATTCAGGAGCA 840
Db   781 ATTGCTCAAACTAAAACCACTACTGACTAAATCAAATAATCCTTCAACTTTTAATTCAGGAGCA 840

Qy   841 ATGCCTGGTGCAACAAACATAGATACGATTCTCAAGTCTCAAGTAAGAAATAAAACA 900
Db   841 ATGCCTGGTGCAACAAACATAGATACGATTCTCAAGTCTCAAGTAAGAAATAAAACA 900

Qy   901 TCTTTCCAATTAGATGAAAATTTGTTATCCAGAAATGGACTGGTTCTGAAGAGAATAAA 960
Db   901 TCTTTCCAATTAGATGAAAATTTGTTATCCAGAAATGGACTGGTTCTGAAGAGAATAAA 960

Qy   961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020
Db   961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAATTTCAAGACTA 1080
Db  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAATTTCAAGACTA 1080

Qy  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGTTGTAGTAGTATTTATATTTTGGTACC 1140
Db  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGTTGTAGTAGTATTTATATTTTGGTACC 1140

Qy  1141 TCTGAATTACCATCATTATGGTACTATTCCCAACTAGATTATCTGATCTAACCGCT 1200
Db  1141 TCTGAATTACCATCATTATGGTACTATTCCCAACTAGATTATCTGATCTAACCGCT 1200

Qy  1201 TTGAATCAAGTTAAAACAGATATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260
Db  1201 TTGAATCAAGTTAAAACAGATATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260

Qy  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320
Db  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320
```

; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: mgc3 gene
US-10-131-591A-2

Query Match 100.0%; Score 3189; DB 15; Length 3189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ATGAATATTTCTAAAAAACTTAAAAGTATACATTGATAAGTGGATTAGCTGTATTTGGA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAATATTTCTAAAAAACTTAAAAGTATACATTGATAAGTGGATTAGCTGTATTTGGA 60

Qy   61 GCTCTTGGTTCTGCAAGCTTTGGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCTCTTGGTTCTGCAAGCTTTGGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120

Qy  121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTGAGACTTGCAGGTCTTGGACAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTGAGACTTGCAGGTCTTGGACAA 180

Qy  181 AATGGTTCGTGTTCAATACAGTTCTTAGAGATGTTGATGAATAACTTTATAACAGCAGCT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AATGGTTCGTGTTCAATACAGTTCTTAGAGATGTTGATGAATAACTTTATAACAGCAGCT 240

Qy  241 AATGGAACAATTATCAAATTAGATAGTTTTACTCAGCAATTGGTAGTACATTA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AATGGAACAATTATCAAATTAGATAGTTTTACTCAGCAATTGGTAGTACATTA 300

Qy  301 GATGATTGTGGTGATACAAAGTAAAACAAATAGTTCAGATTACACAACTAGCAGAAAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GATGATTGTGGTGATACAAAGTAAAACAAATAGTTCAGATTACACAACTAGCAGAAAT 360

Qy  361 AGATTTGATCAAGACAAACAAGAGCATATTATGTCTCTGTTGGTTAATGAAGAGCTAAC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AGATTTGATCAAGACAAACAAGAGCATATTATGTCTCTGTTGGTTAATGAAGAGCTAAC 420

Qy  421 GTTCATTTAAAAGAATTAATACTCAAATAGAATTGGTAATAGAACAACAATTCT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTTCATTTAAAAGAATTAATACTCAAATAGAATTGGTAATAGAACAACAATTCT 480

Qy  481 AAGTTTGTAATTGGTGGTGTGATAATCCAGCTCACGTAATTACTGATGATGGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AAGTTTGTAATTGGTGGTGTGATAATCCAGCTCACGTAATTACTGATGATGGG 540

Qy  541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGAT 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGAT 600

Qy  601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTGGTATACATTCAAAGAAAG 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTGGTATACATTCAAAGAAAG 660

Qy  661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTGATTGTGGAATGGTCAAATAACTAATAACAGATCCT 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTGATTGTGGAATGGTCAAATAACTAATAACAGATCCT 720

Qy  721 ACAAATGCTGATGATGGGATGTTGATTGTGGGAATGGTGGTCAAATAACTAATAACAGATCCT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACAAATGCTGATGATGGGATGTTGATTGTGGGAATGGTGGTCAAATAACTAATAACAGATCCT 780

Qy  781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGGCA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGGCA 840

Qy  841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 900

Qy  901 TCTTTCCAATTAGATGATGAAAATTTGTTTATCCAGAATGACTGGTTCTGAAGAGAATAA 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCTTTCCAATTAGATGATGAAAATTTGTTTATCCAGAATGACTGGTTCTGAAGAGAATAA 960
```

```
Qy 2401 TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460

Qy 2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGAT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGAT 2520

Qy 2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAGAACTTCGTTACAAACCAA 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAGAACTTCGTTACAAACCAA 2580

Qy 2581 AGCTTCTTAAACAGTTAGTTGACTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 AGCTTCTTAAACAGTTAGTTGACTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640

Qy 2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 2700

Qy 2701 TACTTAGATGTGTAAGTATTATGATGCTAAATTAAAGAACAATAATTAGTTAGTAACATTCTCT 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 TACTTAGATGTGTAAGTATTATGATGCTAAATTAAAGAACAATAATTAGTTAGTAACATTCTCT 2760

Qy 2761 TATAACAACTTGGGCGCCTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TATAACAACTTGGGCGCCTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820

Qy 2821 GGTATTCTTCTGCAATTATGATCATCTTAGGATTAGTAGTCCGGTATTCGTATTCCTTTAAGGAGCTCAA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 GGTATTCTTCTGCAATTATGATCATCTTAGGATTAGTAGTCCGGTATTCGTATTCCTTTAAGGAGCTCAA 2880

Qy 2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940

Qy 2941 GCTGTTGGTTCAGTTTACAGAGATTATTACCCCAAACTGCTCAACGTTAAGAGAAAAAACCT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 GCTGTTGGTTCAGTTTACAGAGATTATTACCCCAAACTGCTCAACGTTAAGAGAAAAAACCT 3000

Qy 3001 GCTGCTTAGGTGCTGGTAAATCTCGTGGTAAAGAGAAACCTCGTGCTGCTGCTAAACCTGCT 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 GCTGCTTAGGTGCTGGTAAATCTCGTGGTAAAGAGAAACCTCGTGCTGCTGCTAAACCTGCT 3060

Qy 3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGTAAACCAACTGGGCCTAAA 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGTAAACCAACTGGGCCTAAA 3120

Qy 3121 TCTGGTGCGCGCCTACAAAACCAACCAACTGCTCCTAAGCCAGCTCGTCCAAAAACCAACCGCTCCC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TCTGGTGCGCGCCTACAAAACCAACCAACTGCTCCTAAGCCAGCTCGTCCAAAAACCAACCGCTCCC 3180

Qy 3181 AAAGAATAA 3189
         |||||||||
Db 3181 AAAGAATAA 3189
```

RESULT 2
US-10-131-591A-2
; Sequence 2, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SEQ ID NO 2
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 3189
; TYPE: DNA

```
Qy    961  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCGAACGAAAGATATTGGATTCTTGAC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCGAACGAAAGATATTGGATTCTTGAC  1020

Qy   1021  ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTCAAGACTA   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTCAAGACTA   1080

Qy   1081  TACTTAAACTCAGTAATTCTTATCATTCATTCGTGATAGTAATTTATATTTTTGGTACC   1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TACTTAAACTCAGTAATTCTTATCATTCATTCGTGATAGTAATTTATATTTTTGGTACC   1140

Qy   1141  TCTGAATTACCATCATTATGTACTAATTCCAACTAGATATCTGATCTAACCGCT       1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  TCTGAATTACCATCATTATGTACTAATTCCAACTAGATATCTGATCTAACCGCT       1200

Qy   1201  TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA   1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA   1260

Qy   1261  AACGGAACAACGACAACAGCTGATCACATCTAGTGGTTCAACAGGTGCTGAACAGGAAAT  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  AACGGAACAACGACAACAGCTGATCACATCTAGTGGTTCAACAGGTGCTGAACAGGAAAT  1320

Qy   1321  ACTACTAACACTTCTCAAACACTTCTAATCCTACTTAAATACTTATCGTAGTTTTGGA    1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  ACTACTAACACTTCTCAAACACTTCTAATCCTACTTAAATACTTATCGTAGTTTTGGA    1380

Qy   1381  ATTGATGATAAACCAACTTCTGCAAACAAAATAGATGAAATAAATTGGGCAGATCCTAAC  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  ATTGATGATAAACCAACTTCTGCAAACAAAATAGATGAAATAAATTGGGCAGATCCTAAC  1440

Qy   1441  GTTATTGAAGCAAGAATAATATGCTGATAACATCTAGTGGTATTCAAAATGAAATTCCAATA  1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  GTTATTGAAGCAAGAATAATATGCTGATAACATCTAGTGGTATTCAAAATGAAATTCCAATA  1500

Qy   1501  ACTAATGCAGAGGAAAACTTTATCCGAAACACAATTGGTGTGTTTACTTCAACAGGT    1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  ACTAATGCAGAGGAAAACTTTATCCGAAACACAATTGGTGTGTTTACTTCAACAGGT    1560

Qy   1561  TCAAGAGTAGTTTTAAGAGCTTCTTATATAACGTTCAACGTCCAACTGGAAACTTCCAA   1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  TCAAGAGTAGTTTTAAGAGCTTCTTATATAACGTTCAACGTCCAACTGGAAACTTCCAA   1620

Qy   1621  CCTTTCTTATACGTATTTGGTTATTTAGGATCAACCAACAAACTAGAACAGGAACTTTCTGG   1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  CCTTTCTTATACGTATTTGGTTATTTAGGATCAACCAACAAACTAGAACAGGAACTTTCTGG   1680

Qy   1681  TACGGAACATATAAGCTTTAAACAACAGCGCCTACGACGTATTAGATTCTCCAAGAGTA    1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  TACGGAACATATAAGCTTTAAACAACAGCGCCTACGACGTATTAGATTCTCCAAGAGTA    1740

Qy   1741  GGTACTGAAACCAATCAATTAGAAGAACTTCATTAGAAATAACTTAAAGTTGGGGTTGGTGGATAT  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  GGTACTGAAACCAATCAATTAGAAGAACTTCATTAGAAATAACTTAAAGTTGGGGTTGGTGGATAT  1800

Qy   1801  CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATAAAGAGCACAAGGTGAC    1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATAAAGAGCACAAGGTGAC    1860

Qy   1861  ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAAATCTTATGAGTACATT  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAAATCTTATGAGTACATT  1920

Qy   1921  CAATCAGTTTAGGATTTGATGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA    1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  CAATCAGTTTAGGATTTGATGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA    1980

Qy   1981  AGCTTCTTAAACTCAAATGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA   2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  AGCTTCTTAAACTCAAATGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA   2040

Qy   2041  TACTTAAGATCACAAATTGGATTAGCTAGAGAAACATCTGATTACCAAACCAACCATTC   2100
```

```
Qy   2041  TACTTAAGATCACAAATTGGATTAGCTAGAGAAACATCTGATTACCAAACCAACCATTC   2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  TACTTAAGATCACAAATTGGATTAGCTAGAGAAACATCTGATTACCAAACCAACCATTC   2100

Qy   2101  GGAACAACTCACCAAGTTACTTCAGTATTCAGTATCAGTTCTCATCAGTTCTCATCAATTAAGAAT  2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  GGAACAACTCACCAAGTTACTTCAGTATTCAGTATCAGTTCTCATCAGTTCTCATCAATTAAGAAT  2160

Qy   2161  ATTAGAACAATCTTCCCTGGTAACCAGTATGTGTACTTCTTATTCACAAATGAAAATAT   2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  ATTAGAACAATCTTCCCTGGTAACCAGTATGTGTACTTCTTATTCACAAATGAAAATAAT  2220

Qy   2221  AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA   2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA   2280

Qy   2281  TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTTATTAGAC        2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTTATTAGAC        2340

Qy   2341  AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGT    2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGT    2400

Qy   2401  TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTAATCTGATATTGCTTTT   2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTAATCTGATATTGCTTTT   2460

Qy   2461  GAAGGTTCTGGTGCTGAAGTATACATCTGATTTCTGGGGAACAATCAAATTCAAACCCGAT   2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  GAAGGTTCTGGTGCTGAAGTATACATCTGATTTCTGGGGAACAATCAAATTCAAACCCGAT   2520

Qy   2521  GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA   2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA   2580

Qy   2581  AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA   2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA   2640

Qy   2641  GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA   2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA   2700

Qy   2701  TACTTAGAATGGTAAGTATGATATGCTAAATTAAGAACAATAATTAGTAACAATTCTCT    2760
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  TACTTAGAATGGTAAGTATGATATGCTAAATTAAGAACAATAATTAGTAACAATTCTCT    2760

Qy   2761  TATAACAACTTGGGCGCCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA        2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761  TATAACAACTTGGGCGCCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA        2820

Qy   2821  GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCCTTGATAAGAAACCTGCTCTGCTCAA  2880
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821  GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCCTTGATAAGAAACCTGCTCTGCTCAA  2880

Qy   2881  AGAAAATTACAAGACAAAGGGTTCAAAAACACAATTCAAAAAAGTTGATACCTTGACTGCT   2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881  AGAAAATTACAAGACAAAGGGTTCAAAAACACAATTCAAAAAAGTTGATACCTTGACTGCT   2940

Qy   2941  GCTGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCT    3000
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941  GCTGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCT    3000

Qy   3001  GCTGCTTTAGGTGCTGGTAAAATCTGGTGATAAGAAATCGGTGCTAAACCTGCT        3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001  GCTGCTTTAGGTGCTGGTAAAATCTGGTGATAAGAAATCGGTGCTAAACCTGCT        3060

Qy   3061  GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGTAAACCTGGGCCTAAA     3120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061  GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGTAAACCTGGGCCTAAA     3120

Qy   3121  TCTGGTGCGCCTACAAAACCAACTGCTCCTCAAGCCAGCTGCTCCAAAACCGCTCCC    3180
```

```
Db  3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTCAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3180

Qy  3181 AAAGAATAA 3189
         |||||||||
Db  3181 AAAGAATAA 3189
```

RESULT 3
US-10-125-818-1
; Sequence 1, Application US/10125818
; Publication No. US20020187162A1
GENERAL INFORMATION:
; APPLICANT: University of Connecticut
; APPLICANT: Geary, Steven J.
; APPLICANT: Sekellick, Margaret
; APPLICANT: Marcus, Philip J.
; APPLICANT: Silbart, Lawrence
; TITLE OF INVENTION: Use of a Live Attenuated Mycoplasma gallisepticum Strain as a Vac
; TITLE OF INVENTION: Vector for the Protection of Chickens and Turkeys from Respirato
; FILE REFERENCE: 883923.0074
; CURRENT APPLICATION NUMBER: US/10/125,818
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,569
; PRIOR FILING DATE: 2001-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Mycoplasma gallisepticum
US-10-125-818-1

Query Match   98.6%; Score 3143.4; DB 14; Length 8354;
Best Local Similarity   99.1%; Pred. No. 0;
Matches 3159; Conservative 2; Mismatches 28; Indels 0; Gaps 0;
```

```
Qy     1 ATGAATATTTCTAAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3778 ATGAATATTTCTAAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTGTATTTGGA 3837

Qy    61 GCTCTTGGTTCTGCGAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3838 GCTCTTGGTTCTGCGAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 3897

Qy   121 TTAGTTAATCAAGCAAGAACGCTAGATGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3898 TTAGTTAATCAAGCAAGAACGCTAGATGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA 3957

Qy   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3958 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCT 4017

Qy   241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4018 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 4077

Qy   301 GATGATTGTGTGGATCACAAGTAAAACAAGTAAAACAAATAGTTTCAGATTACACAACTAGT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4078 GATGATTGTGTGGATCACAAGTAAAACAAGTAAAACAAATAGTTTCAGATTACACAACTAGT 4137

Qy   361 AGATTTGATCAAGAACAAGACAAGAGCAGCATATATGCTCGTGTTGGTTAATGATGAAGCTAAC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4138 AGATTTGATCAAGAACAAGACAAGAGCAGCATATATGCTCGTGTTGGTTAATGATGAAGCTAAC 4197

Qy   421 GTTCATTTAAAAAGAATTAATACTAACTAACAAATTGGTAATAGAATTGGTAATAGAAACAACAATTCT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4198 GTTCATTTAAAAAGAATTAATACTAACTAACAAATTGGTAATAGAATTGGTAATAGAAACAACAATTCT 4257

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCAGCTCACGTAATTAGATTTACTGATGATGGG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4258 AAGTTTGTAATTGGTGGTGTTGATAATCAGCTCACGTAATTAGATTTACTGATGATGGG 4317

Qy   541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4318 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 4377

Qy   601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAAACTTATACATTCAAAGAAAG 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4378 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAAACTTATACATTCAAAGAAAG 4437

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4438 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGACCCAGTAGGTAGAGTTAGTGGA 4497

Qy   721 ACAAATGCTGATGATGGGATGTTTGATGTGGGAAATGGTCAAATAACTAATAACTAATAACAGATCCT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4498 ACAAATGCTGATGATGGGATGTTTGATTTGGGAAATGGTCAAATAACTAATAACTAATAACAGATCCT 4557

Qy   781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTCAACTTTAATTCAGGAGCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4558 ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTCAACTTTAATTCAGGAGCA 4617

Qy   841 ATGCTGGTGCAACAATAGATACGAATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4618 ATGCCTGGTGCAACAATAGATACGAATTCTCAATTGAATGTCAAGCATAGAATTAAAACA 4677

Qy   901 TCTTTCCAATTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAA 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4678 TCTTTCCAATTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAA 4737

Qy   961 AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCACGAAAGATATTGGATTCTTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4738 AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCACGAAAGATATTGAATTCTTGAC 4797

Qy  1021 ATACCCGGGACTCCACAAGTTACTTAAAAGAAGATTCAGTTAACGTATTTTTCAAGACTA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4798 ATACCAGGGACTCCACAAGTTACTTAAAAGAAGATTCAGTTAACGTATTTTTCAAGACTA 4857

Qy  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTATTTATTTTTGGTACC 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4858 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTATTTATTTTTGGTACT 4917

Qy  1141 TCTGAATTACCATCATTATGGTACTATATTCCAACTAGATATCTGATCTAACCGCT 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4918 TCTGAATTACCATCATTATGGTACTATATTCCAACTAGATATCTGATCTAACCGCT 4977

Qy  1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACACA 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4978 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 5037

Qy  1261 AACGGAACAACGACAACAGCTGATGACATCTAGTGGTTCAACAGTTATAACAGCAGCT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5038 AACGGAACAACGACAACAGCTGATGACATCTAGTGGTTCAACAGTTATAACAGCAGCT 5097

Qy  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTAAATACTTTAAATATGGTTTATTCGTAGTTTTGGA 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5098 ACTACTAACACTTCTCAAACAGTTCTAATCCTAAATACTTTAAATATGGTTTATTCGTAGTTTTGGA 5157

Qy  1381 ATTGATAGTAAACCAACTTCTGCAAACAAAATAGAAGTAATTGGGCAGATCCTAAC 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5158 ATTGATAGTAAACCAACTTCTGCAAACAAAATAGAAGTAATTGAGCAGATCCTAAT 5217

Qy  1441 GTTATTGAAGCAAGAATATGCTGAATCAGATTAGGATTACAAAATGAAATTCCAATA 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5218 GTTATTGAAGCAAGAATATGCTGAATACAGATTAGGATTACAAAATGAAATTCCAATA 5277

Qy  1501 ACTAAATGCAGGAAACTTATCCGAAACACACAATTGGTGTGGTTTACTTCAACAGT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5278 ACTAATGCAGGAAACTTATCCGAAACACAATTGGTGGTTTGGTTTACTTCAACAGT 5337

Qy  1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTCATCGAACTTCCAACTGGAAGATGGG 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5338 TCAAGAGTAGTTTTAAGACTTTCTTATAACGGTCACTCGAACTTCCAACTGGAAGATGGG 5397

Qy  1621 CCTTTCTTATACGTATTTGGTTATTTAGGATACCAACAAACTAGAACACAGACTTTCTG 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 4
US-10-131-591A-79
; Sequence 79, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
;  APPLICANT: Nippon Zeon Co., Ltd.
;  TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
;  FILE REFERENCE: J209
;  CURRENT APPLICATION NUMBER: US/10/131,591A
;  CURRENT FILING DATE: 2002-08-15
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;  LENGTH: 3129
;  TYPE: DNA
;  ORGANISM: Mycoplasma gallisepticum
;  FEATURE:
;  OTHER INFORMATION: Modified mgc3 gene
US-10-131-591A-79

  Query Match             95.9%;  Score 3057.8;  DB 15;  Length 3129;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches  3083;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;
```

Sequence alignment data (Query/Db) with nucleotide positions from 124 to 6966.

```
Qy  305 ATTGTGGTGATACAAGTAAACAAACAAAATAGTTTCAGATTACACAACTAGCAGAAATAGAT 364
        ||||| |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db  245 ATTTTGCTCGATACAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGAT 304

Qy  365 TTGATCAAAGACAAACAAGAGCATATTATGCTCGTCTGTTGTTAATGATGAAGCTAACGTTC 424
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  305 TTGATCAAAGACAAACAAGAGCATATTATGCTCGTCTGTTGGTTAATGATGAAGCTAACGTTC 364

Qy  425 ATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGT 484
        ||||      |||||||||||||||||||||||||||||||||  ||||||| |||||||
Db  365 ATTTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACCAAAATTCTAAGT 424

Qy  485 TTGGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTA 544
        |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  425 TTGGTAATTGGTGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTA 484

Qy  545 AATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGGC 604
        ||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  485 AATTTCAGTTTACACCAGCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGGC 544

Qy  605 CAATCTTACCTAAGATTTACACCCAGATTGGTATAACTTATCATTCAAAGAAAGATCT 664
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  545 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATCATTCAAAGAAAGATCT 604

Qy  665 TACCAAATGACGTCAACACTGCAGTTGTTCCTTGCCCAGTAGGTAGAGTTAGTGGAACAA 724
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  605 TACCAAATGACGTCAACACTGCAGTTGTTCCTTGCCCAGTAGGTAGAGTTAGTGGAACAA 664

Qy  725 ATGCTGATGATGGAGTGTTGATTGTTGATTTTGGGAATGGTCAAATAACTAACTAACAGATCCTATTG 784
        ||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  665 ATGCTGATGATGGGATGGTTGATTGTTGATTTTGGGAATGGTCAAATAACTAACTAACAGATCCTATTG 724

Qy  785 CTCAAACTAAAACCACTACTAGATATGCTCCAAGCAACGAAAGATATTGGATTCTTCAGGAGCAATGC 844
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  725 CTCAAACTAAAACCACTACTAGATATGCTCCAAGCAACGAAAGATATTGGATTCTTCAGGAGCAATGC 784

Qy  845 CTGGTGCAAACAATAGATAGATACGATTCTCAATTGAAATGTCAAGCATAGAATTAAAACATCTT 904
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  785 CTGGTGCAAACAATAGATAGATACGATTCTCAATTGAAATGTCAAGCATAGAATTAAAACATCTT 844

Qy  905 TCCAATTAGATGAAAAATTGTTTATCCAGAATGACTGGTTCTGAAGAGAATAAAAATA 964
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  845 TCCAATTAGATGAAAAATTGTTTATCCAGAATGACTGGTTCTGAAGAGAATAAAAATACT 904

Qy  965 TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC 1024
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  905 TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC 964

Qy 1025 CCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTATACT 1084
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  965 CCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTATACT 1024

Qy 1085 TAAACTCAGTTAATTCTTATCATCATTGGTGATAGTAGTATTATATTTTGGTACCTCTG 1144
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1025 TAAACTCAGTTAATTCTTATCATCATTGGTGATAGTAGTATTATATTTTGGTACCTCTG 1084

Qy 1145 AATTACCATCATTATGGTACTATCATTATCCAACTAGATATCGATCTAACCGCTTGA 1204
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1085 AATTACCATCATTATGGTACTATCATTATCCAACTAGATATCGATCTAACCGCTTGA 1144

Qy 1205 ATCAAGTTAAAACAGATGATATTGAACTTCAAGCTTCAGTTGTTCAACAGGTACAACAAACG 1264
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1145 ATCAAGTTAAAACAGATGATATTGAACTTCAAGCTTCAGTTGTTCAACAGGTACAACAACACG 1204

Qy 1265 GAACAATCTTCCCTGGTAACCAGTTATGGTCAACAGGTGCTGGAACAGGAAATACTA 1324
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1205 GAACAATCTTCCCTGGTAACCAGTTATGGTCAACAGGTGCTGGAACAGGACAGACTA 1264

Qy 1325 CTAACACTTCTCAAACAGTTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGAATTG 1384
        |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1265 CTCAGATTCTCAAACAGTTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGAATTG 1324
```

```
Qy 1385 ATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAACGTTA 1444
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db 1325 ATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAACGTTA 1384

Qy 1445 TTGAAGCAAGAATATATGCTGATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1504
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1385 TTGAAGCAAGAATATATGCTGATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1444

Qy 1505 ATGCAGGAAACTTATCCGAAACACAATTGGTGGTGTTTTACTTCAACAGGTTCAA 1564
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1445 ATGCAGGAAACTTATCCGAAACACAATTGGTGGTGTTTTACTTCAACAGGTTCAA 1504

Qy 1565 GAGTAGTTTTAGAGCTTCTTATAACGGTGATCAAGTCCAACTGGAAACTTCCAACCTT 1624
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1505 GAGTAGTTTTAGAGCTTCTTATAACGGTGATCAAGTCCAACTGGAAACTTCCAACCTT 1564

Qy 1625 TCTTATACGTATTGGTTATTTAGGATACCAACAACTAGAACAGGAACTTTCTGGTACG 1684
        |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 1565 TCTTATACGTATTTGGTTATTTAGGATACCAACAACTAGAACAGGAACTTTCTGGTACG 1624

Qy 1685 GAACATATAAGCTTTAAACAACAGACCCTTACGACGTATTAGATTCTCCAAGAGTAGGTA 1744
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1625 GAACATATAAGCTTTACAGAACGCCCTTACGACGTATTAGATGTGCGAAGAGTAGGTA 1684

Qy 1745 CTGAAACCAATCAATTAGAAGAACTTCATTAACAACTAACATACCCTGTTATGGGTGATATCTAA 1804
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1685 CTGAAACCAATCAATTAGAAGAACTTCATTAACAACTAACATACCCTGTTATGGGTGATATCTAA 1744

Qy 1805 CTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATAGAGCACAAGGTGACACAC 1864
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1745 CTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATAGAGCACAAGGTGACACACT 1804

Qy 1865 CAGAAAGCCGAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATTCAAT 1924
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1805 CAGAAAGCCGAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATTCAAT 1864

Qy 1925 CAGTTTTTAGGATTTGATGGAATTAGAAATAACTAAAATGTTGGGGTTAAAGCATCAAGCT 1984
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1865 CAGTTTTTAGGATTTGATGGAATTAGAAATAACTAAAATGTTGGGGTTAAAGCATCAAGCT 1924

Qy 1985 TCTTAAACTCAAATAGACCAAATCCAAACGTCTAGAAATGATTGCTGCAACAACATACT 2044
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1925 TCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACATACT 1984

Qy 2045 TAAGATCACAAATTGGATTAGTAGCTAGAGATATTGGATTCTCAACCACCATTCGGAA 2104
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1985 TAAGATCACAAATTGGATTAGTAGCTAGAGATATTGGATTCTCAACCACCATTCGGAA 2044

Qy 2105 CAACTCACCAAGTTATTCAGTATCACCTGGTGATCAGTTGTCATCAATTAAGAATATTA 2164
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2045 CAACTCACCAAGTTATTCAGTATCACCTGGTGATCAGTTGTCATCAATTAAGAATATTA 2104

Qy 2165 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAATAAAT 2224
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2105 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATCAGAAAT 2164

Qy 2225 CTAGTGTTTATACAATTAAGAGATTGACGTAAGAAATTGGTGTAATCTTATTAGACAATT 2284
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2165 CTAGTGTTTATACAATTAAGAGATTGACGTAAGAAATTGGTGTAATCTTATTAGACAATT 2224

Qy 2285 GTCCAACAAGTTTAATTGACGTAATGCTGCTTGGTAATGTTGCATTGTTCTCATCAAACCCTGTTCTC 2344
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2225 GTCCAACAAGTTTAATTGACGTAATGCTGCTTGGTAATGTTGCATTGTTCTCATCAAACCCTGGTTCTC 2284

Qy 2345 CATTCTATACAGTAAATGCTGCTGAACAGGAGAAATACTA 2404
        |||||||||||||||||||||||||||||||||||||||||
Db 2285 CATTCTATACAGTAAATGCTGCTGAACAGGAGAAATACTA 2344

Qy 2405 CTGGATCATATACGTCGTGTAAATAACATTTAATCAGAACTTAATCGATATTGCTTTTGAAG 2464
        ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db 2345 CTGGATCATATACGTCGTGTAAATACATTTAATCAGAACTTAATCGATATTGCTTTTGAAG 2404

Qy 2465 GTTCTGGTGCTAAGTATACATCTGATTTCTCGGGGAACAATCCAATTCAAACCCGATGAGT 2524
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2465 ...
```

Db   2405 GTTCTGGTGCTAAGTAGTACATCTGATTTCTCGGGGAACAATCCAATTCAAACCCGATGAGT 2464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2525 ACTTAATTCAAAATGGGTTCACTAGTCAGTCAGTGGCTAGAGAAACTTCGTTACAAACCAAAGCT 2584
           |||||||||||||||||| |||||||||||||||||||||| | |||||||||||||||||

Db   2465 ACTTAATTCAAAATGGGTTCACTAGTCAGTGGCTAGAGAAACTTCGTTACCACAGCAAAGCT 2524

Qy   2585 TCTTAAACAGTTTAGTTGTTGACTTCACTCCTGCTAATGCTGGTACTCACCGTGTAGTGG 2644
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2525 TCTTAAACAGTTTAGTTGTTGACTTCACTCCTGCTAATGCTGGTACTCACCGTGTAGTGG 2584

Qy   2645 TTGATCCTGATGGTGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAATACT 2704
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2585 TTGATCCTGATGGTCAGCTGACAACCAAAACCTACCTCTAAAAGTTCAGATCCAATACT 2644

Qy   2705 TAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTTAGTAACATTCTTATA 2764
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||

Db   2645 TAGATGGTAAGTATTATGATGCTAAATTAAGAACAATAATTTAGTAACATTCTTTATA 2704

Qy   2765 ACAACTTTGGCGCCCTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTA 2824
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||

Db   2705 ACAACTTTGGCGCCCTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTA 2764

Qy   2825 TTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTAAGAGCTCAAAGAA 2884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2765 TTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTAAGAGCTCAAAGAA 2824

Qy   2885 AATTACAAGACAACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCTGCTG 2944
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2825 AATTACAAGACAACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCTGCTG 2884

Qy   2945 TTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCTGCTG 3004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2885 TTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCTGCTG 2944

Qy   3005 CTTTAGGTGCTGCGGTAAATCTGGTGATAAGAAAACTGGTGATAAGAAAACCTGCTGCTC 3064
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2945 CTTTAGGTGCTGCGGTAAATCTGGTGATAAGAAAACTGGTGATAAGAAAACCTGCTGCTC 3004

Qy   3065 CAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAGCTCACCAAACCAACTGCGCCT 3124
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   3005 CAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAGCTCACCAAACCAACTGCGCCT 3064

Qy   3125 GTGCGGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGGCTCCCAAAG 3184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   3065 GTGCGGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGGCTCCCAAAG 3124

Qy   3185 AATAA 3189
           |||||

Db   3125 AATAA 3129

RESULT 5
US-10-282-122A-27316
; Sequence 27316, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27316
; LENGTH: 3159
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-10-282-122A-27316

Query Match   3.9%; Score 124.6; DB 13; Length 3159;
Best Local Similarity 63.5%; Pred. No. 1.6e-15;
Matches 190; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy   2770 TTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
           |||||| ||||||||||||| ||| ||||||||||||||||||||||||||||

Db   2797 TTTGCAGCACTCCCTGCATGGGTGATCCCTGTATCAGTTGGGATCTTG 2856

Qy   2830 GCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAAAGAAAATTA 2889
           |||||||||||||||||||||||| |||||||||||||||||||

Db   2857 TTTATCTTGTTAGTCTTAGGACTTGGGATCCCAATGTACACGGTAAGAAAAACTC 2916

Qy   2890 CAAGACAAAGGGTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCTGTTGGT 2949
           ||||| ||||||||||||||

Db   2917 CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAACTGCTGTCGGT 2976

Qy   2950 TCAGTTTACAGAAGATTATTACCCAAACTGCTAACGTTGTGTGGTGAAAAAAGCACCTGCTTTA 3009
           |||||||

Db   2977 AGTGTGTACAAAAGATTATTACCCAAACTGGTGATAAGAAAAAAGCACCTAGTAGTGCATTG 3036

Qy   3010 GGTGCTGGTAAATCTGGTGATAAGAAAATCGGTAAACCTGCTAAACCTGCTCTCCAGC 3068
           |||||||

Db   3037 AAAGCTGCTAATCCTAGTGTTAAAAAACTGCTGCTTTTTAAAACCACCTGTTCAACC 3095

RESULT 6
US-10-205-220-1
; Sequence 1, Application US/10205220
; Publication No. US20030170663A1
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmer
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1D1
; CURRENT APPLICATION NUMBER: US/10/205,220
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 08/545,528
; PRIOR FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 1
;   LENGTH: 580073
;   TYPE: DNA
;   ORGANISM: Mycoplasma genitalium
US-10-205-220-1

Query Match    3.9%;    Score 124.6;  DB 15;  Length 580073;
Best Local Similarity  63.5%;  Pred. No. 2.9e-14;
Matches  190;  Conservative  0;  Mismatches 109;  Indels  0;  Gaps  0;

Qy   2770  TTTGGGCCTTACCTTACCTCATGGTGCCTACAGCAATTGGTAGTACATTAGGTAGTATTCTT  2829
Db  228703  TTTGCAGCACTCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG  228762
Qy   2830  GCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAAAGAAATTA   2889
Db  228763  TTTATCTTGTTTAGTCTTTAGGACTGGGATTGGGATCCCAATGTACAGGGTAAGAAACTC  228822
Qy   2890  CAAGACAAAGGGTTCAAAACAACATTCAAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT  2949
Db  228823  CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATCACTCACCAACTGCTGTCGGT  228882
Qy   2950  TCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAGGAGAAAAACCTGCTGCTTTA  3009
Db  228883  AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGGTGAAAAAGCACCTAGTGCATTG  228942
Qy   3010  GGTGCTGGTAAATCTGGTGATAAGAACCTGCTGCTGCTAAACCTGCTGCTCCAGC    3068
Db  228943  AAAGCTGCTAATCCTAGTGTTAAAAAACCTGCTGCTTTTTAAAAACCACCTGTTCAACC  229001


RESULT 7
US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
;   APPLICANT: ZWIEBEL, LAURENCE J.
;   TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
;   FILE REFERENCE: N8289
;   CURRENT APPLICATION NUMBER: US/10/094,240
;   CURRENT FILING DATE: 2001-03-08
;   PRIOR APPLICATION NUMBER: 10/056,405
;   PRIOR FILING DATE: 2002-01-24
;   PRIOR APPLICATION NUMBER: 60/264,649
;   PRIOR FILING DATE: 2001-01-26
;   NUMBER OF SEQ ID NOS: 27
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4985
;   TYPE: DNA
;   ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match    2.0%;   Score 62.6;  DB 15;  Length 4985;
Best Local Similarity  45.5%;  Pred. No. 0.02;
Matches  333;  Conservative  0;  Mismatches  396;  Indels  3;  Gaps  3;

Qy    253  ATCAAATTAGATAGTTTCTAAAACCATTATATGGTTTAGATCTAAGTGATGATTGTGGT   312
Db   3070  ATAAAAAAAAACAATAATAAACAATAATAAGAATAACAATAACAATAGAGATTGATCA   3011
Qy    313  GGATACAAAGTAAAACAAATAGTTTCA-GATTACAACAACTAGCAGAAATAGATTTGATCA  371
Db   3010  CAACAACAACAACAATAATAATACAATAACAATAGCAATAATAATAATAAGAATAATAA  2951
Qy    372  AAGACAAAACAAGAGCACATATTATGCTCGTTGTTGTTAATGATGAAGCTAAGTTTGTAA  430
Db   2950  TAACAATAGCAATAATAGTAATTAATAAGAATAATAATACAATAATAATAATAATAA    2891
Qy    431  AAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAACAATTCTAAGTTTGTAA  490
Db   2890  AAAAAAATAATTAATAATAAGAATAGTAATAATAATAATAAAATATATTAAAAAATAATATTA  2831
```

```
Qy    491  TTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGGACTAAATTTA        550
Db   2830  ATAATAATAATAATATAATATAAGAATAATAATAATAATATAGATTATAATAAT ATAATAATAATA  2771
Qy    551  ATTTTACAAACCAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCT  610
Db   2770  ATAATAATAATAATCATAATAAAATTATTATTATGATAATAATGTATAATAATGAAT  2711
Qy    611  TAC-CTAAAGATTTACCCCAGATTGGTATAACTTATCATTCAAAGAAGATCTTACCA   669
Db   2710  AACAATAATAATAATAATTAATAAAATTAATAATTAATAATAATAATAATAATAAT  2651
Qy    670  AATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTTCTTCAGTTGGAACAAATGCT  729
Db   2650  TATTATTGTTATAATAATAAGAGTAAATAATAATAATAATCATAATTATAAT        2591
Qy    730  GATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAA  789
Db   2590  AATAATAATAATAATAATAATAATAATAATACAGATAATAATAATAATAATAAA     2531
Qy    790  ACTAAAACCACTACTGATAATCAAAAATCCTTCAAACTTTAATTCAGTAATATTAATATAT  849
Db   2530  AGTAATAATTATAATATAGTAATATAATAATAATAATAATAATATTAATAATATAAT   2471
Qy    850  GCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAAACATCTTTCCAA  909
Db   2470  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT     2411
Qy    910  TTAGATGAAAAATTGTTTATCCAGAATGGACTGGTTCTGAAGAATGAATAAAAATATTACA  969
Db   2410  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT   2351
Qy    970  AGATTAGCTACT   981
Db   2350  AAATTCTGTATT  2339


RESULT 8
US-10-056-405-10/c
; Sequence 10, Application US/10055405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
;   APPLICANT: ZWIEBEL, LAURENCE J.
;   TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
;   TITLE OF INVENTION: USE THEREOF
;   FILE REFERENCE: N7841
;   CURRENT APPLICATION NUMBER: US/10/056,405
;   CURRENT FILING DATE: 2002-01-24
;   PRIOR APPLICATION NUMBER: 60/264,649
;   PRIOR FILING DATE: 2001-01-26
;   NUMBER OF SEQ ID NOS: 23
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4985
;   TYPE: DNA
;   ORGANISM: Anopheles gambiae
US-10-056-405-10

Query Match    2.0%;   Score 62.6;  DB 15;  Length 4985;
Best Local Similarity  45.5%;  Pred. No. 0.02;
Matches  333;  Conservative  0;  Mismatches  396;  Indels  3;  Gaps  3;

Qy    253  ATCAAATTAGATAGTTTACTAAAACCATTATATGGTTTAGATCTAAGTGATGATTGTGGT   312
Db   3070  ATAAAAAAAAACAATAATAAACAATAATAAGAATAACAATAACAACAATAATAAGAA   3011
Qy    313  GGATACAAAGTAAAACAAATAGTTTCA-GATTACAACAACTAGCAGAAATAGATTTGATCA  371
Db   3010  CAACAACAACAATAATAATACAATAATAATAGCAATAATAATAATAAGAATAATAA    2951
Qy    372  AAGACAAAACAAGAGCACATATTATGCTCGTTGTTGTTAATGATGAAGCTAAGTTCAT-TTAA  430
Db   2950  TAACAATAGCAATAATAGTAATTAATAAGAATAATAATACAATAATAATAATAATAA     2891
```

```
Qy   256  AAAGAGTAGTTTACTAACCATTATATATGTTTAGATCTAAGTGATGATGTGGTGGA  315
          |||  ||| ||  ||   ||| || |  |||||| ||| ||| ||| |
Db  5053  ATAATCTATACATACAATAAATATATTATTCAACTTTAAAAAAAAAAAACTTAATCTA  4994

Qy   316  TACAAAGTAAAACAAATAGTTCAGATTACCAACTAGCAGAAAATAGATTTGATCAAAGA  375
          |||| |  ||||||  ||||| |||||  |  |||| |||||||| ||| ||
Db  4993  TACTACACAAAAAAAAAACCTTAAAAACATTATCAAAAATATTATACAATAACAAAAA  4934

Qy   376  CAAACAAGAGCATATTATGCTCGTCGTTGGTGTTGGTTAATGATGAAGTTCATTTAAAAAGA  435
          ||| || |||  ||||| |  |  |   ||||||  ||| ||||| ||||||
Db  4933  AATACATACTATATATAATTCCACTTCTAAAAAATAACTACCTAAAATAATCAAATTATAAAAAAC  4874

Qy   436  ATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTGTAATTGGT  495
          || |  || || | ||| ||| |||| ||| |||||  || ||  |||
Db  4873  AAAATAATATACATAAATTCAAAAACCTAAAAAAAAAAAAATAAAAATTATTTATTAAT  4814

Qy   496  GGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAATTTAATTTT  555
          ||| |  |||| ||| | |||||||  |  ||  || ||||  | ||  |||| |||
Db  4813  AAATAATAATAATATTATACAAAAAATATAAAACATAATTAATAAATTATACCACACA  4754

Qy   556  ACAAACCAAACTCAAGGTGAAATTGTAATGACTTCATTTTAGATGGCGGCCAATCTTACCT  615
          || |||| | | ||| ||| ||||||| ||| |||| || ||| || ||||||
Db  4753  TACAAATAATCAAAATAAATTTATATATAATTTTATTATCATATTTATTATATATTTACCA  4694

Qy   616  AAAGATTTACACCCAGATTGGTATAACTTATACTTATATCAAAGAAAGATCTTACCCAAATGAC  675
          ||| |||| || ||| ||||| |||| ||||  ||| |  |||| |||  ||| ||
Db  4693  TAATACAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAACTTAACGATTTACATAACTTA  4634

Qy   676  GTCAACACTGCA  687
          |  ||| ||
Db  4633  ACCAATACCTCA  4622
```

Query Match   1.9%; Score 60.2; DB 13; Length 3489;
Best Local Similarity 61.6%; Pred. No. 0.052;
Matches 114; Conservative 0; Mismatches 68; Indels 3; Gaps 1;

```
Qy  2998  CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT  3057
          |||||||||  ||| |||| |||||| |||| || |||||||||||||||
Db  2236  CATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCTCATCCT  2177

Qy  3058  GCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTGCGTGCTCATCCTGGGC--  3115
          |||||||||  |||| ||| ||| |  || || ||||| |||||||||||||
Db  2176  GCTGCTGCTGCTCATCCTGCTGCTCATCCTCTCCTCATCCTGCTGCTGCTCCT  2117

Qy  3116  -CTAAATCTGGTGCGGCCTACAAAACCAACTGCTCCTAGCCAGCTCCAAAACCAACC  3174
           |   |||| ||||| | ||| ||||||||||||||||| |||||||||||
Db  2116  GCTCATCCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTCATCCTGCT  2057
```

```
Qy   431  AAAGAGTTAATACTACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTGTAA  490
          ||||  ||| ||   || ||  ||||| |||| |||| || ||||| |  |||
Db  2890  AAAAAAATAATTAATTATATAATAAGAATAGTAATAATAATATAATTAAAAATAATATTA  2831

Qy   491  TTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATTTACTGATGATGGGACTAAATTA  550
          || | | | |||| |   |  | | ||||| || ||  || ||| |||  ||
Db  2830  ATAATAATAATAATAATATAATAAGGATAATAATATAATATAATATAATATAATATAATATA  2771

Qy   551  ATTTTACAAACCAAACTCAAGGTGAAATTGTAATGACTTCATTTTTAGATGATGCGCCAATCT  610
          || || || | |||  |||| | ||| ||| ||| |||  |||| |||  ||||||
Db  2770  ATAATAATAATAATAATCATAATAAAATTATTATGATAATAATGTAATAATAATAATGAAT  2711

Qy   611  TAC-CTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAGATCTTACCA  669
          ||  | || | || || ||| |||| |||| || || |  ||||| ||| |||||
Db  2710  AACAATAATAATAATTAATAAATAAATTCAAAAAACCTAAAAAAAAAAATAAATAATAAT  2651

Qy   670  AATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTGAGTTAGTGGAACAAATGCT  729
          |||  | |||| ||| || |  | |  || |  || ||  ||||  || | || |||
Db  2650  TATTATTGTTATAATAATAATAATAATAATAATATAATAATAATCATAATTATAAT  2591

Qy   730  GATGATGGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAA  789
          |||  |  || |  | |||| ||| |||| |||| || || ||||| || ||
Db  2590  AATAATAATAATAATAATAATAATAATTATAATAATAATAATAATAATAATAATAAA  2531

Qy   790  ACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGAGAGCAATGCCTGGT  849
          || |||||| || |||| |||  ||||  | || || || || |||||  |||||||
Db  2530  AGTAATAATAATAATAATAATAGTAATAATAATAATAATAATAATAATAATAATAAT  2471

Qy   850  GCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATGAATTAAAAACATCTTTCCAA  909
          ||||| || || || ||||| ||||| || |  ||| ||| ||||| || || |||
Db  2470  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT  2411

Qy   910  TTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAAATATTACA  969
          || |  |||||| || ||  || | ||| |  ||  | |||||| |||||  ||||||
Db  2410  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT  2351

Qy   970  AGATTAGCTACT  981
          | ||| | ||
Db  2350  AAATTCTGTATT  2339
```

Query Match   1.9%; Score 60.8; DB 15; Length 8771;
Best Local Similarity 46.3%; Pred. No. 0.065;
Matches 200; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

```
Qy    3175 GCTCCC 3179
           ||||||
Db    2056 GCTGC 2052
```

RESULT 11
US-10-294-804-1/c
; Sequence 1, Application US10294804
; Publication No. US20030133948A1
; GENERAL INFORMATION:
;   APPLICANT: Robertson, Erle S.
;   APPLICANT: Cotter, Murray A.
;   TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
;   TITLE OF INVENTION: to Genomic Host DNA
;   FILE REFERENCE: UM-03778
;   CURRENT APPLICATION NUMBER: US/10/294,804
;   CURRENT FILING DATE: 2002-11-14
;   PRIOR APPLICATION NUMBER: US/09/410,399
;   PRIOR FILING DATE: 1999-10-01
;   NUMBER OF SEQ ID NOS: 6
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3489
;   TYPE: DNA
;   ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-10-294-804-1

  Query Match   1.9%; Score 60.2; DB 15; Length 3489;
  Best Local Similarity   61.6%; Pred. No. 0.052;
  Matches 114; Conservative   0; Mismatches 68; Indels   3; Gaps   1;

```
Qy    2998 CCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT 3057
           ||||||| ||| ||||| ||||||||| | ||||||||||||||||||||||||||
Db    2236 CATCCTGCTGCTGCTCATCCTGCTGCTCGCTGCTCATCCTGCTGCTGCTCATCCT 2177

Qy    3058 GCTGCTCCAGCTAAACCATCTGCACCAAAGCTAGCTCCACCCAGCTAAACCAACTGGGC-- 3115
           ||||||| ||||||||||||| ||||||||| |||||||||||||||||||||||||
Db    2176 GCTGCTGCTGCTCATCCTGCTGCTGTGCTGCTCCTGCTCATCCTGCTGCTGCTCCT 2117

Qy    3116 -CTAAAATCTGGTGCGCCTACAAAAACAACTGCTCCTAGCCAGCTGCTCCAAAACCAACC 3174
            |||||  |||||||||||  ||||||||||||| |||||||||||||||||||||||
Db    2116 GCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCATCCTGCT 2057

Qy    3175 GCTGCC 3179
           ||||||
Db    2056 GCTGCC 2052
```

RESULT 12
US-09-790-988-1
; Sequence 1, Application US09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
;   APPLICANT: SHIGENOBU, SHUJI
;   APPLICANT: WATANABE, HIDEMI
;   APPLICANT: HATTORI, MASAHIRA
;   APPLICANT: SAKAKI, YOSHIYUKI
;   TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;   FILE REFERENCE: 081356/0159
;   CURRENT APPLICATION NUMBER: US/09/790,988
;   CURRENT FILING DATE: 2001-02-23
;   PRIOR APPLICATION NUMBER: JP2000-107160
;   PRIOR FILING DATE: 2000-04-07
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match   1.8%; Score 58.4; DB 9; Length 640681;
  Best Local Similarity   47.3%; Pred. No. 2.3;
  Matches 209; Conservative   0; Mismatches 231; Indels   2; Gaps   1;

```
Qy     241 AATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATGGTTTAGATCTAAGT 300
           |||||   ||||| ||||||||| |||| ||  ||||||| || |||  |||| |||
Db  525047 ATTAAAAAAAATAAACACAGAAATTTAATACATACCATTCTCAAAGTAATTTTATT 525106

Qy     301 GATGATTGTGTGGATACAAGTAAACTAAAACAAATAGTTTCAGATTACAACTAGCAGAAAT 360
           |||||| |||||| |||| |||||| ||||| ||||| ||| |||||| ||||||||
Db  525107 TTAAAAAAGAGAAGAAAAATAATATTATAAAACTATTATCTCAAAAAAGAATAAAAAAA 525166

Qy     361 AGATTTGATCAAAGACAAGAAGAGACATATATGCTCTGTGGTTAATGATGAAGCTAAC 420
           ||||| ||| |||| ||||| ||| |||||||| ||  ||||||  |||||||| ||
Db  525167 GCAATTGTTAAAATAGATCCTTCAATCTATAAAAATATAACAAATCAAGAAGCACAA 525226

Qy     421 GTTCATTTAAAAGATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT 480
           ||| ||||||| |||||| |||| | ||||| ||| |||| ||||||| | ||||| |
Db  525227 ATTTATTTCAAAAAAAATCAAGATAATTTTTATATTCCAGAAAAATTTAAAATTAATTTT 525286

Qy     481 AAGTTTGTAATTGGTGTGTTGATAATCCAGCTCACGTAATTTATGATGATTACTGATGATGGG 540
           ||||||||  ||||| || |||||| |||| ||| |||  ||| | ||||| |||||||
Db  525287 GTTGAATTAAAAACTGATAACTTAAAATACATTGCGAAATAAAAATAAAGAGATTTATGATTGG 525346

Qy     541 ACTAAATTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTCATTTTAGAT 600
           ||||||||||||| |||||| |||| ||||| |||| || | |  ||| || ||||
Db  525347 TATATAAGAAATATTAC--ACAATACTCAACTAAAGAAAAAAGAGATATAGTATTATTC 525404

Qy     601 GCGGCCAATCTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG 660
           || |||||||  ||||||||||| || |||||||||||| | |||||||||||||||||
Db  525405 AAGTCAAAATAAACAACAGCCATATGCATTATTATCTAGATTGCATAAATACACCAGAAG 525464

Qy     661 ATCTTACCAAATGACGTCAACA 682
           |||||  |||||  || ||||
Db  525465 ATTTTTCAAAAATAGCTCAAGA 525486
```

RESULT 13
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US2003018627TA1
; GENERAL INFORMATION:
;   APPLICANT: Epigenomics AG
;   TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
;   FILE REFERENCE: E01/1208/WO
;   CURRENT APPLICATION NUMBER: US/10/312,841
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3294164)
US-10-312-841-1

  Query Match   1.7%; Score 53.6; DB 15; Length 3673778;
  Best Local Similarity   47.1%; Pred. No. 59;
  Matches 196; Conservative   0; Mismatches 219; Indels 1; Gaps   1;

```
Qy     801 TACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAACATGCCTGGTGCAAACAATAG 860
           ||||||  |||||||||||| |||||||| |||| |||  || ||||||| |||||||||
Db   98780 TAATTATCCAATTATTATTCCATCACCATTAATAAAAAAACTCTTCCTTCATTAAAATTA 98721

Qy     861 ATACGAATTCTCAATTGAATGTCAAGCATGACATAGAATTAAAAACATCTTCCAATTAGATGAAAA 920
           |||| |||| |||| ||||| |||  |||| ||| |||||||| ||||| |||  ||||| |||
Db   98720 CTTTTATACTTTTATAAAACATTAAAAATCAATTAAAACATATTTAATTTACTCATTTTAATTAAATTACTC 98661

Qy     921 ATTGTTTATCCAGAATGACTGGTTCTGAACGAAGATTAAAACTTTCACCATATTCTCAATTTCATCACA 980
           |||||||| |||||| |||| |||| |||| ||| |||||| || | |||| ||||| ||||
Db   98660 ATCTTCTACATTTACAATTAAAATTAAAAACATTTCAGATTCATAAATACATATTCTTCAATTTCATCACA 98601
```

```
Qy    981  TGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGACATACCCGGGACTCCACAGT 1040
                |||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  98600  AATATAATTTATAATATATTATAAAATATTTTTCTCAATCTACGACTTATCTTCTCATT 98541

Qy   1041  TACTTTAAAAGAAGAGATTCAGTTAACGTATTTCAAGACTATACTTAAACTCAGTTAATTC 1100
                ||||||||||| | ||||||||||||||||||||||||||||| |||
Db  98540  TTCTCAATATTATATCTTTTAAAATACTTAAATTTTAACTAAACCCAATTTA-TC 98482

Qy   1101  TTTATCATTCATTGGTGATAGTATTTATTATTTTGGTACCCTCTGAATTACCATCATTATG 1160
                ||| | ||||||||  |||||||||||||||||||
Db  98481  TCTCACTTTTTTTTCATCTATAAAATCATATTTTTTATATTTTTATATAAAAATCTTTTACC 98422

Qy   1161  GTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCTTTGAATCAAGTTAAAA 1216
                ||| ||||||||  |||| | ||
Db  98421  TAACTACTCAATATCACAAAAATTTTCTCCTATATTTTCTTTAAAAATTTATAA 98366
```

RESULT 14
US-10-373-667-3
; Sequence 3, Application US/10373667
; Publication No. US20030235841A1
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P. W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/10/373,667
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US/09/181,585
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 3
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA
; OTHER INFORMATION: comprising exons E, C, and A
US-10-373-667-3

Query Match   1.7%; Score 53.2; DB 16; Length 1037;
Best Local Similarity   50.8%; Pred. No. 0.77;
Matches 127; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```
Qy   2930  CCTTGACTGCTGCTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTA 2989
                ||||||   ||||||||
Db    605  CCTGGGTCCTTCATGTTAGAAAACCTGGCTTACTACTACTACTACTACTA 664

Qy   2990  AGAAAAAACCTGCTACTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTG 3049
                |||||||||||||||| |||||||||||||||||||
Db    665  CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724

Qy   3050  CTAAACTGCTGCTCCAGCTAAACCATCTGCACCAAAAGTAGCTCACCAGCTAAACCAA 3109
                ||||||||||||
Db    725  CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784

Qy   3110  CTGGGGCCTAAATCTGGTGCGCCTACAAAACCAACTGCTGCTGCTGCTTCCAAAAC 3169
                |||||||
Db    785  CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844

Qy   3170  CAACCGCTCC 3179
                |||||
Db    845  CTGCTGCTGC 854
```

RESULT 15
US-10-373-667-1
; Sequence 1, Application US/10373667
; Publication No. US20030235841A1
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P. W.
; APPLICANT: Koob, Michael

; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/10/373,667
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US/09/181,585
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 1
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-373-667-1

Query Match   1.7%; Score 53.2; DB 16; Length 1159;
Best Local Similarity   50.8%; Pred. No. 0.82;
Matches 127; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```
Qy   2930  CCTTGACTGCTGCTGTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTA 2989
                ||||||   ||||||||
Db    417  CCTGGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA 476

Qy   2990  AGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTG 3049
                |||  ||||||||| |||||||||||||||||||
Db    477  CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536

Qy   3050  CTAAACTGCTGCTCCAGCTAAACCATCTGCACCAAAAGTAGCTCACCAGCTAAACCAA 3109
                |||||||||||
Db    537  CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596

Qy   3110  CTGGGGCCTAAATCTGGTGCGCCTACAAAACCAACTGCTGCTGCTGCTCCTAAGCCAGCTGCTCCAAAAC 3169
                ||||||||
Db    597  CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656

Qy   3170  CAACCGCTCC 3179
                |||||
Db    657  CTGCTGCTGC 666
```

Search completed: May 6, 2004, 00:19:49
Job time : 1270.74 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2004, 10:50:12 ; Search time 7289.65 Seconds
                 (without alignments)
                 13063.796 Million cell updates/sec

Title: US-09-901-572A-2
Perfect score: 3189
Sequence: 1 atgaatatttctaaaaact........caaccgctcccaagaataa 3189

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfum:*
                 16: em_estom:*
                 17: em_gss_hum:*
                 18: em_gss_inv:*
                 19: em_gss_pln:*
                 20: em_gss_vrt:*
                 21: em_gss_fun:*
                 22: em_gss_mam:*
                 23: em_gss_mus:*
                 24: em_gss_pro:*
                 25: em_gss_rod:*
                 26: em_gss_phg:*
                 27: em_gss_vrl:*
                 28: gb_gss1:*
                 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Accession	Description
1	c	73.6	2.3	1896	29	CG753083	CG753083	P048-1-C0
2	c	70.8	2.2	1101	29	CNS0039G	AL063921	Drosophil
3		67	2.1	1200	13	BX437758	BX437758	Drosophil
4		66.4	2.1	1101	29	CNS00EVL	AL069706	Drosophil
5		65.8	2.1	844	29	BX139987	BX139987	Danio rer
6	c	65.2	2.0	1101	29	CNS0039G	AL063921	Drosophil
7		64.6	2.0	829	29	BX173672	BX173672	Danio rer
8		64.2	2.0	773	29	CNS01VTG	AL169549	Tetraodon
9	c	63.8	2.0	1200	13	BX447454	BX447454	F.rubripe
10		61.8	1.9	470	29	FR0018463	AL011359	F.rubripe
11	c	61.8	1.9	480	28	BZ643398	BZ643398	OGAOW57TC
12	c	61.8	1.9	843	28	BZ643413	BZ643413	OGAOW57TM
13	c	61.8	1.9	1225	29	CNS0161D	AL106171	Drosophil
14	c	61.6	1.9	735	29	CNS04NSM	AL299119	Tetraodon
15		60.6	1.9	641	28	AQ946120	AQ946120	Sheared D
16	c	60.2	1.9	1101	29	CNS00FVE	AL071298	Drosophil
17		60	1.9	941	13	BU126589	BU126589	603003315
18	c	60	1.9	1056	29	BX415058	BX415058	F.rubripe
19		59.8	1.9	450	29	FR0025683	AL018519	F.rubripe
20		59.8	1.9	498	12	BJ396861	BJ396861	BJ396861
21		59.8	1.9	575	28	BZ780846	BZ780846	ii20h11.g
22		59.8	1.9	615	28	BH766948	BH766948	BMBAC345F
23		59.8	1.9	994	13	BX414650	BX414650	BX414650
24	c	59.2	1.9	619	29	FR0006944	Z90754	F.rubripes
25	c	59.2	1.9	1201	13	BX458623	BX458623	BX458623
26		59		317	12	BJ328366	BJ328366	BJ328366
27		58.8	1.8	619	29	PR0013713	AL004959	P.rubripe
28		58.8	1.8	759	29	CNS06QXV	AL411257	T7 end of
29	c	58.8	1.8	766	14	CF348501	CF348501	AGENCOURT
30		58.8	1.8	813	29	BX143161	BX143161	Danio rer
31	c	58.8	1.8	836	28	BH491109	BH491109	BOGGY47TR
32		58.8	1.8	1007	29	CNS06X9S	AL419462	T3 end of
33		58.4	1.8	1126	13	BX446391	BX446391	
34		58.2	1.8	332	13	C92788	C92788	C92788 Dict
35		58.2	1.8	641	12	BM181884	BM181884	fv51bll.y
36	c	58.2	1.8	1350	29	CG744271	CG744271	P036-4-C0
37		58	1.8	257	29	FR0026883	AL019712	F.rubripe
38		58	1.8	939	29	CG900918	CG900918	ZMMBBb051
39	c	57.4	1.8	1101	29	CNS0100X	AL098379	Drosophil
40	c	57.4	1.8	1101	29	CNS00BSI	AL069797	Drosophil
41		57.2	1.8	427	28	BZ422321	BZ422321	ids2c10.b
42		57.2	1.8	649	12	BM160056	BM160056	EST562579
43	c	57.2	1.8	1101	29	CNS001FB	AL060732	Drosophil
44	c	56.8	1.8	639	29	CNS038CX	AL232458	Tetraodon
45	c	56.8	1.8	989	29	CNS02HA4	AL197365	Tetraodon

ALIGNMENTS

RESULT 1
CG753083/c
LOCUS   CG753083   1896 bp   DNA   linear   GSS 24-OCT-2003
DEFINITION P048-1-C01.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG753083
VERSION   CG753083.1 GI:37977199
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1896)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE   An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de

```
FEATURES   Class: BAC ends.
source     Location/Qualifiers
           1..1896
           /organism="Pristionchus pacificus"
           /mol_type="genomic DNA"
           /strain="California"
           /db_xref="taxon:54126"
           /clone_lib="Ppa EcoRI BAC Library"
           /note="The library was generated by a partial digest of
           the genomic DNA with EcoRI and cloning into the BAC
           vector."

ORIGIN

Query Match        2.3%;  Score 73.6;  DB 29;  Length 1896;
Best Local Similarity  42.8%;  Pred. No. 7.5e-05;
Matches  512;  Conservative  0;  Mismatches  675;  Indels  9;  Gaps  4;

Qy    315 ATACAAGTAAAACAAATAGTTCAGATTCACAACTAGCACAACATGAGATTGATCAAAG 374
          |  | ||| ||||||   |||    |  ||||||   |||| ||   ||  |||
Db   1786 AAAAAATTAAAAAATAAAATAAAATAAAATAATANNAATTTAAATAAATATATATATA 1727

Qy    375 ACAAACAAGAGCATATTATGCTCGTCTGTTGGTTAGATATATTTNATAAATANANTAAA-A 433
          |||| || |    | | || | | ||   ||  |     |||||
Db   1726 AATATTATTAATTAATTATTNATAATATTNATAATATTAAAATTATATTATATANAATA 1667

Qy    434 GAATTAATACTAACTCAAATAGATAGAGTTGGTAATAGAAACAACAATTCTAAGTTGTAATTG 493
          ||||| | |  ||| ||  |  ||| | ||   ||| || |  | | |   | ||| |||
Db   1666 TAATTAAAAATAATAAATATATAATAATTANNAAAAAAAAAAATATTTTNAATTTATATATA 1607

Qy    494 GTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAATTTTAATT 553
          |  | ||| ||||   | |   |||| |  |  | |  | || ||| ||| | ||||||
Db   1606 TAATAATAATTATTTAAAATATNATTAAATTAATAATTATTTTATATATTTTTTTATTTTTAT 1547

Qy    554 TTACAAACCAAACTCAAGTGAAATTGTTAATGACTTCATTTTAGATGCGGCCAATCTTAC 613
          ||    | ||||  |  || |||||| || |   | |  |||  |||
Db   1546 TTTTAATATATAATATTTTTTAATATTTTATTTAAAAATAATATATAATAAAATTAT 1487

Qy    614 CTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCTT--ACCAAA 671
          |||| ||| || |   |||| |||||   |  | |  ||  | |  | |||   |||||
Db   1486 ATTTTTTTTTATTTTTTTTTTATNAAATATTATTATATATAATAAAATTATNAAAATAA 1427

Qy    672 TGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTAGTGGAACAAATGCTGA 731
          | | ||||| |  || | |||||    |||| |  || ||| ||| ||| | |||| ||
Db   1426 AAAAAATTATAAANTAAATTTATTTTATTTATTAAAATATTAAAATATTATTNANNAA 1367

Qy    732 TGATGGGATCGTTGATTGTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAAAC 791
          |||| | || || |  | | |   |  || || || | ||||| ||| || | || |||
Db   1366 TATAAAATATATAAAAATATNATAAAATAAATTAAATTATATTTTTTTAAATTTAT 1307

Qy    792 TAAAACCACTACTGATAATCAAATCCTTCAACTTTAATTCAGGAGCAATGCCTGGTGC 851
          |||| | | | ||| ||| |  |  ||  |||| | ||||| || | |||| ||| ||
Db   1306 TAATTAAATTATAAAATAAAAATAAAAATTTATATAAA---TAATAATAATATNA 1251

Qy    852 AAACAATAGATACGATTCTCAATGAATGTCAAGCATAGAATTAAAACATCTTCCAATT 911
          |||  ||| || | ||| |  |||| || |||| ||  |||| || |||||  ||||
Db   1250 AAAAAATAAAAAAAAAAAAATAATAATAATAATAATAATATAAAAAATAAATAAT 1191

Qy    912 AGATGAAAAATTGTTATCCAGAATGGACTGGTTCTGAAGAAGAATAAAAATATTACAAG 971
          | ||| ||| || || | | ||||| || |||| |||||||| | |||||||| |||
Db   1190 TTAATAAAATTAATTATTATAAATAATAATAATATAAATAAAAAAATAAAAAATAAAAAAA 1131

Qy    972 ATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATATGGATTCTTGCACATACCCGGGAC 1031
          || | | | ||| |||| | |||| || ||||| ||||| |||| ||||||| |||
Db   1130 ATAAAATTATATAAATTAAAAATAAAATNATTATTTNATTTATTAATTTAAAAAATATAATAT 1071

Qy   1032 TCCAAGTTTACTTTAAAGAAGAGATTCAGTTAACGTAATTTCAAGACTATACTTAAACTC 1091
          |  || |||| |||||| || || |||| ||| |||| | |||| | || |||| | ||
Db   1070 TTATTTATTTATATTATATAATTANTATTATTTNTTTATATATATTAATAAAAAATTA 1011

Qy   1092 AGTTAATTCTTTATCATTCATTGGTGATAGTATTATTTATTTATTTTGGTACCTCGAATTACC 1151
          ||||| || |||| | ||| || | || || ||||| ||||| || |||| |||| ||||||
Db   1010 TAATTAAATTTTTAAATAAAATAAATAAATANATATTTTATTTATTAATTTTATTAAAAATATAT 951
```

```
Qy   1152 ATCATTATGGTACTATTCCAACTAGATTAATCGATCTGATCTAACCGCTTTGAATCAAGT 1211
          | |||| || | | |||  ||| | |||||    |  || |||| | ||| ||| |||||
Db    950 TAAAAAATAAATTATTTAATTAATTAAAAAANAATTAAAATTATTATTATATATATA 891

Qy   1212 TAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACAACGGAACAAC 1271
          |||| | | || | |||| | ||||| | ||||| ||| ||| |  |  || |||| |
Db    890 AAAAAAAATAAAATAAATTAAAAAAAAAAATANAAATAAAATATATATATAATTTATA 831

Qy   1272 GACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAATACTACT--AAC 1329
          || |||| | ||| | || | ||| |||||| || || |||| || |||| ||   |||
Db    830 TATAATTTATNATNATAANAAAAAATTATAAAATAAAATATAATAATTTTTATATTTAAAT 771

Qy   1330 ACTTCTCAAACAGTTCTAATCCTACTTTAAATACTTAAATACTTATCGTAGTTTGGAATTGATAGT 1389
          | |||| ||| | || | || ||| ||||| |   |||| ||  | |  ||||| ||||| || ||
Db    770 ATTATTAAATTATATAATAATAATAATAATAATTTATAATAATAAAATAAAAAATAA 711

Qy   1390 AAACCAACTTCTGCAAACACAAAATAGATGAAACTAATTGGGCAGATCCTAACGTTATTGAA 1449
          ||| ||| | | |||| | |||||| | || |||||| |  ||| ||||||| ||| |||
Db    710 AATTAAAAATATATAATATATATATTTTTTTTTATTTATATTTTATAATAATAATATAATAA 651

Qy   1450 GCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTAA 1505
          || | ||| |||| | || || || ||| ||| ||||| ||| ||| ||||| |||
Db    650 TTATNNAAATATAATAATTATTATAAAATATAATTATATAAAATATTATATAAAATA 595
```

```
RESULT 2
CNS0039G/c
LOCUS       CNS0039G   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TBT3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly) genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org the BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES    Location/Qualifiers
source      1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR08K10"
            /clone_lib="RPCI-98"
            /note="end : TBT3"

ORIGIN

Query Match        2.2%;  Score 70.8;  DB 29;  Length 1101;
Best Local Similarity  16.3%;  Pred. No. 0.00028;
Matches  94;  Conservative  255;  Mismatches  227;  Indels  0;  Gaps  0;
```

```
Qy    7  ATTTCTAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGAGCTCTT  66
          | |::| |||:: :|| ::: ||:  :|::: :|:: |::: |::: :
Db  1051  AGTWTWTATWWWWWWATWDTWDKWWWWATAAKTDTAWTWWRTAWRADWAGRDRGAGKR  992

Qy   67  GGTTCTGCAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAATTAGTT 126
          ||: :|  |  |  ||| :: :|::: ||: :| : :|: |:::::: :::::
Db  991  DRDAATDADGAGRRDGGRKRKDGGRKRKDGDDDKKGGKKKKAAKAKWATKWWDDWDWDKD 932

Qy  127  AATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAAAATGGT 186
          :  :::::|| ||:  |: :| |:| ||| :|::|| :::|:  |: ||:::
Db  931  WKWDGAKDRKADDDDGAGDKDDDGKGKDADDDTDGTKDDDKDKWDDWDKAKGTWGDATW  872

Qy  187  TCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCTAATGGA 246
          |:| |::||: :::: ||| | |::|:| |:: ::|:|| :|: :|:::|::| ::|:
Db  871  AWAATDWWWGWGWADADWWWTWDLAADDWWADDRWDAWWAKWDDDAWAWGARTADRRDWGDRA  812

Qy  247  ACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTAAGTGATGAT 306
          :: :||::| :|:|: | : ::|::|::|: | :| |::|| :|: ::
Db  811  GKRRGGARKRRDRKRADDKRDAADDRDDAAWTWWTWTTTTRDTDDWKWKTDTWTRWAADRT 752

Qy  307  TGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGATTT 366
          |::|:|::|| || :: ||::::| ::|| :|| |:::|:: |:::::|:|::|:||
Db  751  WDRDDDDRDRAGTAGRKWRRITWKRRKWKRRDTRWDDADADDTARDDRRRGDDGADAGKG  692

Qy  367  GATCAAAGACAAACAAGAGCATATTATGCTCCTGTTGGTTAATGATGAAGCTAACGTTCAT 426
          | ::||:|| :::::|::|:| | | |:::|::|::||:| |::|::| ::| :
Db  691  KKTGRKRRRDRRDRATWDRDTDAWWADAAWWTTTDTDTDWDKDRDRRRKGARRRRRTTARAAW 632

Qy  427  TTAAAAAGAATTAAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTT 486
          ||: | :: ||::: |: ::: | ::||: :: :||::| ::: :::: |::||||
Db  631  DWWTWKAWDWAKWDWKTRADRWDRWAADTWTDARKADRDWAKARAWRARRDRARARARADR  572

Qy  487  GTAATTGGTGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGGACTAAA 546
          |::|:|:| |:||| ||| :::| : :: :|:::|::|::|:::|||
Db  571  RWTTKGKTTTATWTTWAARAWAWAWAWATTTATWTTTTWTTTTTTTTTWTTAA  512

Qy  547  TTTAATTTACAAACCAAACTCAAGGTGAAATTGTT 582
          ::::|||:|| ||: ||: ||:|| ::
Db  511  WWWAAWWTATWAAWTAAAWAAAAAWAAATTTTTTT  476

RESULT 3
BX437758
LOCUS        BX437758        1200 bp    mRNA     linear     EST 15-MAY-2003
DEFINITION   BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01
             5-PRIME, mRNA sequence.
ACCESSION    BX437758
VERSION      BX437758.1   GI:30773605
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1200)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0CAP008CA01QP1.
FEATURES             Location/Qualifiers
     source          1..1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP008YB01"
                     /tissue_type="THYMUS"
```

```
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match             2.1%;  Score 67;  DB 13;  Length 1200;
Best Local Similarity  30.9%;  Pred. No. 0.0018;
Matches  283;  Mismatches 440;  Conservative 181;  Indels 13;  Gaps  4;
```

```
Qy   13  AAAAAACTTAAAAGTYATACATTGATAGGTGGATTAGCTGTATTGGAGCTCTTGGTTCT  72
          ||||||||||||||  ::: :|||| :
Db  266  AAAAAAAAAAAAAAAWWWAAAATWWWAAARATWDGATKADDATTGGGGGDWTTTWWWTWTAA  325

Qy   73  GCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAATTAGTTAATCAA 132
          |  :::| ||| ||||||  ::||||||| :|:|:: :| |: :
Db  326  RRRDKTTTTTTTTTWAAAAARATAKRGGRGGADWTAWWTWTTTTATWWWWDDDDAK  385

Qy  133  GCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAAAATGGTTCGTTG 192
          |:| :|| :|| | ||||||||||| :::: | :: ||: : :::|::|:|::
Db  386  AWADAGAATWTTTTTTWWRAGRGARKRDGTTDTRAADWMWTWAAAWWWDKRKTTTTT  445

Qy  193  TTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCTAATGGAACAATT 252
          |||||| ||||||||  -----TTTTTTTWAKAAAAAAARAAATWTWWRAAAAAAWWTW 500
Db  446  TTKKRRAAAAATTTTT-----TTTTTTTWAKAAAAAAARAAATWTWWRAAAAAAWWTW  500

Qy  253  ATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTAAGTGATGATGGT 312
          :: :|||:||| ::|||||||  ||:|:|| || :||::|| | ||
Db  501  WWAARGRARGATTTTTTTTTAAWWWGGGAGRTAWWTWTTWTTTTTTTTTTTTAWWWATWT  560

Qy  313  GGATACAAAGTAAAACAAATAG-TTTCAGATTACACAACTAGCAGAAATAGATTTGATCA 371
          |:| |||:|| ::||:: |::  ::::::: :|::::||::: |:::|::|||
Db  561  ATATTTTTWTWANAAAAASAGAWKAIWTTTTKKTTATWWKRAGGATWTTTTTTTTWTW  620

Qy  372  AAGACAAACAAGAGCATATTATGCTCCTGTTGGTAATGATGAAGCTAACGTTCATTTAAA 431
          || ::|||:::|::|:| | | |:||::|::||:| |::| ::::||: ||::||||
Db  621  AARGDWKWTWTTTTWTTTTTWTTTTTTTWWTGTTTWWWRDGARRTAKATAKWWTWTWWT  680

Qy  432  AAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTTGTAAT  491
          ||: || | | |: :||: ::|: |:: ::: :|| :  |:|: : ||:||||
Db  681  AWAWAGGAARKRRRATWTWTWTWWAAWAWWAGAAAARGAGAAGRAAWTTTTTTTATKAGARRG 740

Qy  492  TGGTGGTTGATAATCCAGCTCACGTAATTACTGATGATGGGACTAAATTTAA  551
          ||::| :|||||||:| || | ::::|::||||||::| ||::|::||||
Db  741  ATWTTWAWWTWTATRARGADTTTTTTWTATAWWTATTTTWWAWAAGATDKAAAAAWWW  800

Qy  552  TTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCTT 611
          ::::|| || ||: || || :: ::::| ||:::: :: :::||| :|| :|:| ||:
Db  801  WTTTWAAAAAATTWWWAGARAAAKATWTWAAAAAAAAAAAAWRARWAAATATATWTTT  860

Qy  612  ACCTA-AAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCTACCAA 670
          |::| |||:|| |::: |:| ||:|||||| :|||:| ::::|||: |||| :|:::||
Db  861  ATATATAAKARAAAKAARAAAAAAATAARARRWKGAAAAAAATAAAAAWWTTWTATAAA 920

Qy  671  ATGACGTCAACACTGCAGTTGTTCCTTGGCCACGTAGTAGTTGAACAAAATGCTG  730
          ||:|:||||||| ||:||| || :||:|::  |||::|:||| ||:: :|||:|:
Db  921  AAAWATTWAWWRAAAAWWWDTTTTATWTAWTAWTATAGAWRAAATAADATWRAA  980

Qy  731  ATGATGGGATGGTTGATTGTTGATTGTGGGAATGCTCAAATAACTAACAGATCCTATTGCTCAA 790
          |:|:|| ||: :||| |||| ||:||::|  |||| || ||:|| | :||: |::||::|||
Db  981  AAAWWTTTTTWTWWARAAAWWDTTTTATWTATWTAWTATAGAWRAATARDATWRAA 1040

Qy  791  CT-----AAAACCACTACTGATAATCAAATCCTTCAACTTTAATTCAGGAGCAATGC 844
          ||     |::: |:||||| || :: : :|| :: |:::: :||::|::||: |
Db  1041 TTATWADAARRAAAGAWATAKDAGAGATAAAWTTTTTTTTTTTWTAAAAAGAAAAA 1100

Qy  845  CTGGTGCAAACAATAGATACGAATTCAATTGAATGTCAAGCATAGAATTAAAAACATCTT  904
          ||:|||||||| ||:| : :||||:|| |:|::|:||||:||:| |||||: |:|:||
Db  1101 TWTTARAAAAAAWATAAAAAWATAAAAAAAAAAAAAAAGAAAAAKADAWWAAGAAAAATAAA 1160

Qy  905  TCCAATTAGATGAAAA  921
```

```
Db  1161 AARAAAAAAGWRAARAW 1177
          ||||      :  :||:|:
```

RESULT 4
CNS00EVL
LOCUS   CNS00EVL   1101 bp   DNA   linear   GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION   AL069706.1 GI:4943949
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

Query Match   2.1%; Score 66.4; DB 29; Length 1101;
Best Local Similarity 35.1%; Pred. No. 0.0024;
Matches 224; Conservative 104; Mismatches 300; Indels 10; Gaps 3;

Qy   4 AATATTTCTAAAAAAAACTTAAAAGTTATACATTGATAGGTGATTAGCTGTATTTGGAGCT 63
            : :|| |||:|| ||| |: || ||::| :|:: :: |||
Db   471 WAATTTWWAAAWAAAWTTATWAATWAATWAAAAWAWWWATTTTTWWWWWTWATTWTTTWAWW 530

Qy   64 CTTGGTTCTGCGAAGCTTTGGCTTAAGCAATCAGATAAGAGATAACGATAACACGCAATTA 123
            :| |||| ||:| || :: ||| :
Db   531 TWTAWTAAAAAAAAAAWATAATTTAAAWWAATAWATAWWATTAAWRAATTAAWAAWTATATTAAATW 590

Qy   124 GTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACACAAAT 183
            ||| || | | :| |:|| |:|::   ::|
Db   591 TATAAATWTWATTAATTAAAAAAATATTTTTWATTAAAATTTTTWATTAAATTTAATTTAWT 650

Qy   184 GGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATACTTTATAACGCAGCTAAT 243
            :|: |: |::||| | :: |::| : ||::||||||:
Db   651 TATTAAATAAWTTATTWWWTAATTTWAATAAATTTWAAAAAAAAAAAAAAWATWA 710

Qy   244 GGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTAAGTGAT 303
             ::|| ||||: |:   ::||: |: |: ||: | |::
Db   711 AWAATWATAWATAAWTTAAAAWAATAAAAWWAATWAWWATATWWATATATWWTTW 770
```

RESULT 5
BX139987
LOCUS   BX139987   844 bp   DNA   linear   GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-99E7, genomic survey sequence.
ACCESSION   BX139987
VERSION   BX139987.1 GI:27971314
KEYWORDS   GSS.
SOURCE   Danio rerio (zebrafish)
ORGANISM   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 844)
AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
TITLE   Direct Submission
JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT   This sequence was generated from the T7 end of BAC 99E7. 99E7 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES   Location/Qualifiers
     source   1..844
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-99E7"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

Query Match   2.1%; Score 65.8; DB 29; Length 844;
Best Local Similarity 45.2%; Pred. No. 0.0031;
Matches 241; Conservative   0; Mismatches 292; Indels 0; Gaps 0;

Qy   47 TAGCTGTATTGGAGCTCTTGGTTCTGCAAGCTTTGGCTTAAGCAATCAGATAAGAGTA 106
            |||| :: ||:   ::||:| | :| ||:| ||
Db   312 TTGCAGTCATTATTGCTTTATACTTTGATCATTTTGTCATTAATAATAACTAATAATA 371

Qy   107 ACGATAACACGCAATTAGTTAATCAGACAAGAACGCTAGATGCTAATAATATCTGTTAGACTTG 166
            | | :|||| || |: | || || ||:| |: :||||:|
Db   372 ATGATAATAATAATAATGATAATAATTGATAATAATAATTGATAATAATAATAATAATA 431

Qy   167 CAGGTCTTGGACAAAATGGTTCGTTGTTGTTCATCAGTTCTTCATAGATCTAAGTGAT 226
            ||:||| :|:|| |:| ::|| |: || :| || || |: | ::|
Db   432 ATAATGATAATAATAATATTGATAATAATTGATAATAATAATGATAATAATGATAATAATA 491
```

Qy    304 GATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGA 363
            |::  ::  |:|: :||:|:|    |   : ||:|    :    : :||::|:|
Db    771 TAWWNATWWAWWWTATAWATAWATAWAATAWAAWAWAWAWATAAATAWATAWATW--AAAWAWAWA 828

Qy    364 TTTGATCAAAGACAAACAAGAGACATTATATGCTCTGTTGGTTAATGATGAAGCTAACGTT 423
            : :||      |:|: ||    | :    |||::  : : :: :|: |::|:|   |: :|
Db    829 TAWWATWATATAWWAAATAWAWWAAAAAAAATWTAATAWATAWAAWAAAAAAATAWAWTTWW 888

Qy    424 CATTTAAAAGAATAATAATACTCAAAATAGAATTGGTAATAGAAACAACAACAATTCTAAG 483
            |  :|   ::||:|     |::|   |:  ||| : :::      :|||  :| ::|
Db    889 TWTTTTWWAWWATAAWAWTAAAAWATAWAAAWAAAWAAA----AAAAAAAATAWAWWWTWATA 943

Qy    484 TTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTCATGATGGGACT 543
            ||::|    : : ||:|| :|     |::|     : :|||   : |::|::|||: |
Db    944 TTWTTRATTAAAWTWTWATWWATTWWAWTTWWAWTWTWTATATWTWTATATWTWWT---AWT 1000

Qy    544 AAATTTAATTTTACAAACCAAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCG 603
            :  | :| |:     |  :|    | |::|||       | :: :|| ||  : :|
Db    1001 AWATAATATTTATTAWWTATATTTTWAWAAWTATATTATATWWTAWATATAWWWTAWATATAWAWWAA 1060

Qy    604 CCAATCTTACCTAAAGATTTACACCCAGATTGGTATAA 641
            ||||: || :|| : ||:||||||    || |::| ||
Db    1061 TTAWTTATATATWTWAAWATWWAAATAWWAAAAWAAWATAW 1098

```
Qy  324  AAAACAAATAGTTTCAGATTATCAACAACTAGCAGAAATAGATTTTGATCAAAGACAAACAAG  383
             ||:|| |||:: |||| |  ::  :|  ::|  :  ||  |||  ||||
Db  419  AAWRAAATAATTWAAAWAWAAAAAAATTWAAAWAWAAAAAWTAAWTTTAWTWAAAAAAA  478

Qy  384  AGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGTTCATTTAAAAGAATTAATAC  443
             ||:| |||||| |::| ||  :|  ||| : :|    ||| |||::
Db  479  AAAATTWTTTTTTTTWATAWTTWTAWWTTWWWTTAAAWAAWAAAWAAAWA  538

Qy  444  TAACTCAAATAGAATTGGTAATAGAAACAACAATCCTAAGTTTGTAATTGGTGGTGTTGA  503
             |||| :  :||    ||||:  ||  :  ||  |  || : :|||:
Db  539  TAAATWTWTW-WTTTTWAAWATAAAMCMAAWYTHTYTYTYTYTYTYTYTMMTWHYHTM  597

Qy  504  TAATCCAGCTCACGTAATTAGATTACTGATGATGGGACTAAAATTTAATTTACAAACCA  563
             :  :    :||  :  ||   |:||  :| :    |||||:|
Db  598  YTHAWAHTTWYHWYHTYAMWHWMTWHWTMWAWWHWTTYTAYYYYTCMYYHMWHHA  657

Qy  564  AACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCTTACCTAAAGATTT  623
             :||: |   ::||  ::||  ||    |:||  |||  :
Db  658  HAHAAAWWTTHTWWTHAYHWATYHYYYYMYCAMMCMCTHTCHHCYCYYYHHYTAHHTHTHH  717

Qy  624  ACACCCAGATTGGTATAACTTATACATTCAAAGAAGAAGATCTTACCAAATGACGTCAACAC  683
             :||: | :  :  :  || :      |:::    ||   |  :
Db  718  WYAHYYMWYYMWAYYWMYCTACTYHYHHHHHHYEHAYHTTWYAWAHAMWMWHHAHYAAAA  777

Qy  684  TGCAGTTGTTCCTTGGCCAGTAGAGGTAGAGTTAGTGAGTGAACAAATGCTGATGATGGGATGTT  743
             ||:  |||  :| |     |   : :  |  :  | :|  |
Db  778  WAAWAITHHYAHTTHYMHHTYMHYHYMYTCCYMCTYHCWHYHTAYTCWTWHHWWMTWT  837

Qy  744  TGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAAACTAAAACCACTAC  803
             : |   : ||:  ::|||  :   :| | |  | ||  :| :|
Db  838  HWYHHTWWHHTTHWAWWHTHTWCWWHHATTWTWATHCWACMTMHWHWWMEMHHHHMAC  897

Qy  804  TGATATCAAAATCCTCAACTTTAATTCAGGAGGCAATGCCTGGTGTGCAAACAATAGATA  863
             : ::||  :: :: : : :  |||  |  ||| :  ||
Db  898  HAHHHTHMCMCHHHMGHTCHHHHTMYHMTCHWMWMEHWHWWMATWMTMTMWMMCC  957

Qy  864  CGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTTCCAATTAGATGAAAAATT  923
             :::||:: :     : :  |::|  |||:|   |  :|:::|
Db  958  MMHHHCHMYHMHMHYMYCCHYYCTCTHATTHYHYMCTCYHYCTWHTYWYWWAWTAHA  1017

Qy  924  TGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAAATATTACAAGATTAGCTACTGG  983
             ::::|:   :  ::  :  :| |  |||||  |  |  |||
Db  1018  MTTATWWWMMEHWWAEHWATWWWWWWATAWWACTCHHWTWYHTHCTWYYHHTYEHMWWAWW  1077

Qy  984  AAGTTTGCCAAGCAAC  999
             |:::: ::|: ::|:
Db  1078  MAWHWHHMYAHYHWAH  1093
```

RESULT 7
BX173672/c
LOCUS   BX173672   829 bp   DNA   linear   GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-150M6, genomic survey sequence.
ACCESSION   BX173672
VERSION   BX173672.1 GI:28005377
KEYWORDS   GSS.
SOURCE   Danio rerio (zebrafish)
  ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 829)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT   This sequence was generated from the SP6 end of BAC 150M6. 150M6 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES   Location/Qualifiers

```
Qy  227  TTATAACAGCAGCTAATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATG  286
             |||| ||||         ||        |||||||||           |||||
Db  492  ATAATATTGATGATTATGATGATAATATATAATAATAATATATATAATATGATAATA  551

Qy  287  GTTTAGAATCTAACTGATCGATGATTGTGTGGATCAAAGTAAAAACAAATAGTTTCAGATTACA  346
             |||||||||               ||||||||||||            ||||||
Db  552  ATAATGATAATAATATAATAATGATAATATATAATATGATGATAATAATAATGATGATA  611

Qy  347  CAACTAGCAGAAATAGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCGTTGTTA  406
             |||||       |||                  |||||||
Db  612  ATAATGATAATAATAACAATAATAATAATGATGAATAATAATAATGATGATAATA  671

Qy  407  ATGATGAAGCTAACGTTCATTTAAAAGAATTAAAATACTCAAATAGAATTAGATTGGTAATA  466
             ||||||||               ||||||||||||
Db  672  ATAATAATAATAATGATGATAATAATAATGATACTAATAATGACAATAATAATAATG  731

Qy  467  GAAACAACAAATCTAAGTTTGTGTAATTGGTGTGTGATAATCCAGCTCACGTGTAATTAGAT  526
             ||||||||               |||||||            ||||||
Db  732  ACAATAATAATAATATAATAATGATAATAATAATGATAATACTAATAATGATGATAATAATG  791

Qy  527  TTACTGATGATGGGACTAAAATTTACAAAACCAAACTCAAGGTGAAATT  579
             |||||||               |||||||
Db  792  ATAATAATGATAATAATAATATAATGATGATAATAATAATGATGATAT  844
```

RESULT 6
CNS0039G
LOCUS   CNS0039G   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION   AL063921.1 GI:4941778
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library,
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

Query Match   2.0%; Score 65.2; DB 29; Length 1101;
Best Local Similarity 19.1%; Pred. No. 0.0042;
Matches 129; Conservative 264; Mismatches 282; Indels 1; Gaps 1;

```
source          1..829
                /organism="Danio rerio"
                /mol_type="genomic DNA"
                /db_xref="taxon:7955"
                /clone="DKEY-15OM6"
                /tissue_type="Testis"
                /note="vector pindigoBAC-536"

ORIGIN

Query Match       2.0%;     Score 64.6;   DB 29;   Length 829;
Best Local Similarity   45.5%;   Pred. No. 0.0056;
Matches 229;   Conservative   0;   Mismatches 274;   Indels   0;   Gaps   0;

Qy    85  TTTAAGCAATCAGATAAGAGTAACGATAACACGCAATAGTTAATCAAGCAAGAACGCTA 144
              ||||||||| |  || |  ||||  ||| ||| |||||||||| |||||||||
Db   669  TTGAAACATTTGAAGAAAAAAACGTAATAATAATAATAATAATAATGATGAT 610

Qy   145  GATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAAAATGGTTCGTTGTTCAATACAGTT 204
              ||||||||||||||||||| ||   ||||||||||||||||||||||||||
Db   609  GATGATGATGATGATAATAATAATAATAGTAATAATAATAATAATAATAGAAAATAAT 550

Qy   205  CTTAGAGATGTTGATGATGATAACTTTATAACAGCAGCTAATGGAACAATTATCAAATTAGAT 264
               ||| ||| |||||   ||||     |||||||  |||||| ||||||||||||||
Db   549  AATAATGATGATGATGATGATAATAATAATAATAGTAATAATAATAATAATATAT 490

Qy   265  AGTTTTACTAAAACCATTATATGGTTTAGACTCAAGTGATGATTGTGGTGGATCACAAAGTA 324
              ||||||       |||   |||||||||||| |||||||||||||||||
Db   489  AGAAATAATAATAATGATGATGATGATGATAATAATAATAGTAATAATAAT 430

Qy   325  AAACAAAATAGTTCAGATTACACAACTAGCAGAGAAATAGATTTGATCAAAGACAAACAAGA 384
               |||||||  ||| || |||| |||    |||||||||||||||| |
Db   429  AATAATAATAATAATAATAATAAAGATAATAATAATAATAGAAAATATAATAACAAT 370

Qy   385  GCATATTATGCTCGTGTTGGTTAATGATGAAGCTAACGTTCATTTAAAAGAATTAATACT 444
              |||||||||||||||  ||||||||||||||||||
Db   369  GATGATGATGATAATAATATTAATATAGTAATAATAATAATAATAAT 310

Qy   445  AACTTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTTGTAATTGGTGGTGTTGAT 504
              |||    |||||  |||||||||   |||| |||||||||||
Db   309  AATAATAGTATATAATATAATAATAATAATGATGAAATAAT 250

Qy   505  AATCCAGCTCACGTAATTAGAGATTTACTGATGATGGGACTAAATTTAATTTTACAAACCAA 564
              |||||||||    |||||||||  |||||||||||||||||  || |
Db   249  GATAATAATAATAATAATAATAGTAATAATAATAATAGAATAATAATAATGATGAT 190

Qy   565  ACTCAAGGTGAAATTGTTAATGA 587
              |||||||  ||||||
Db   189  GATGATGATGATAATAATAATAA 167
```

```
RESULT 8
CNS01VTG
LOCUS       CNS01VTG       773 bp       DNA       linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            199C24 of library G from Tetraodon nigroviridis genomic survey
            sequence.
ACCESSION   AL169549
VERSION     AL169549.1   GI:7807606
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
```

```
                  PUBMED     10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
                  PUBMED     10899143
REFERENCE   3   (bases 1 to 773)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..773
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="199C24"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG199B12LP1-end : T7"
ORIGIN

Query Match       2.0%;     Score 64.2;   DB 29;   Length 773;
Best Local Similarity   43.9%;   Pred. No. 0.0068;
Matches 162;   Conservative   22;   Mismatches 185;   Indels   0;   Gaps   0;

Qy  1240  ACTGATAACGGTACAACAACAACAAACGGAACAACGACAACAGCTGATACATCTAGTGGTTCA 1299
              ||||| ||||| |  ||||   |||||||||||| ||  ||  ||||  |||||
Db   165  ACTGCTACTACWACWACTGCTACWACTAGTGCWACWCWGCTACTGCWACTACTGCTACT 224

Qy  1300  ACAGGTGCTGGAACAGGAAATACTACTACTACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTTA 1359
              || ||||||||| ||||||||||||||||||||  ||| ||||||||||||||||
Db   225  ACTAGTGCTACTACTGCTACTACTGCTACTACTWAGTGCTACTACWAGTGCTACTACTGCT 284

Qy  1360  AATACTTATCGTAGTTTTGGAATTGATGAGTAAACCAACTTCTGCAAACAAAATAGATGAA 1419
              |||||||| ||    |||||||| || |||  |||   || ||||| |||||||
Db   285  ACTACWACTGCWACTACTGCWACTCWACTACTRCWACTACTACTACTACTACTACT 344

Qy  1420  ACTAATTGGGCAGATCCTCAACGTTATTGAAGCAAGAATAATGCTGAATACAGATTAGGT 1479
              ||||| |||| |||||||||   | |  || ||| ||| ||| ||| |||||||||||
Db   345  ACTACTRCTACTACTGCWACCAGCATACATCTGCTACTACTACATGCTACAT 404

Qy  1480  ATTCAAAATGAAATTCCAATAACTAATGCAGGAAACTTATCCGAAACAAATTGGTGGT 1539
              ||| ||  ||| ||||||||  | |||  |  ||   |||   :|| ||||
Db   405  ACTGCAACTACWGCAACTGCWACTRCTACAGCWACTRCTACACWACAACAACTACWRCW 464

Qy  1540  GTTGGTTTTACTTCAACAGGTTCAAGAGTAGTTTAAGAGCTTCTTATAACGGTGATCAA 1599
              |||    || |||||||||     ||||||||  ||
Db   465  ACWRCAACTACTGCTACAACTACTACTRCWACTGCTACTACTACTACTACTACTRCT 524

Qy  1600  CGTCCAACT 1608
              |||||||||
Db   525  ACTACTACT 533
```

```
RESULT 9
BX447454
LOCUS       BX447454       1200 bp       mRNA       linear   EST 22-MAY-2003
DEFINITION  BX447454 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP003YH19
            5-PRIME, mRNA sequence.
ACCESSION   BX447454
VERSION     BX447454.1   GI:31018746
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

SOURCE   Takifugu rubripes (Fugu rubripes)
ORGANISM   Takifugu rubripes
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
           Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
           Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE   1 (bases 1 to 470)
AUTHORS   Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y.,
           Williams,G. and Brenner,S.
TITLE   Direct Submission
JOURNAL   Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
           Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmp.mrc.ac.uk
COMMENT   Vector: pBluescript II KS
           V.type: phagemid
           PRIMER: KS
           DESCR:
           One pass dye-terminator sequencing of cosmid cloned genomic
           sequence.
FEATURES   Location/Qualifiers
     source   1..470
           /organism="Takifugu rubripes"
           /mol_type="genomic DNA"
           /db_xref="taxon:31033"
           /clone="016B10aC6"
           /clone_lib="cosmid 016B10"
ORIGIN

Query Match   1.9%;   Score 61.8;   DB 29;   Length 470;
Best Local Similarity   51.5%;   Pred. No. 0.021;   Indels 0;   Gaps   0;
Matches 135;   Conservative   0;   Mismatches 127;   Indels   0;

```
Qy  2920 AAAGTTGATACCTTGACTGCTGCTGTTGGTTCAGTTTACAAGAAGAGATTATTACCCAAACT 2979
              ||| ||||||  ||||||||||||||||
Db   345 AATGCTGCTACTACTGCTGCTGCTGCTACTGCTGCTACTGCTGCTACTACTGCTACTGCTGCT 286

Qy  2980 GCTAACGTTAAGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCT 3039
              ||  ||||||||||||||  ||||||||||||||||||||||
Db   285 ACTACTGCTGCTACTGCTGCTGCTCCAGTGAACCATCTGCACCAAAAGCTAGCTCACCA 226

Qy  3040 GCTGCTGCTGAAACTGCTGCTCCAGTGAACCATCTGCACCAAAAGCTAGCTCACCA 3099
              |||||||||||  |||||||||||||||||||
Db   225 GCTGCTNCTNCTNCTGCTGCTACTACTGCTGCTACTGCTGCTACTACTGCTGCTACT 166

Qy  3100 GCTAAACCAACTGGGCCTAAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCT 3159
              |||||||||||||||||||||||||||||||||||||||||
Db   165 GCTACTGCTGCTGCTACTACTGCTCTACTGCTCTACTGCTGCTGCTCTGCTCTGCTACTACT 106

Qy  3160 GCTCCAAAACCAACCGCTCCCA 3181
              |||||||||||||||||
Db   105 ACTGCTACTACTGCTGCTGCTA 84
```

RESULT 11
BZ643398
LOCUS   BZ643398   480 bp   DNA   linear   GSS 29-JAN-2003
DEFINITION OGAOW57TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0111J17,
           genomic survey sequence.
ACCESSION   BZ643398
VERSION   BZ643398.1   GI:28104877
KEYWORDS   GSS.
SOURCE   Zea mays
ORGANISM   Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 480)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
           Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
           Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE   Consortium for Maize Genomics
JOURNAL   Unpublished (2002)
COMMENT   Other_GSS: OGAOW57TM
           Contact: Cathy Whitelaw

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1200)
AUTHORS   Li,W.B., Gruber,C.J., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of
           Invitrogen. This sequence belongs to sequence cluster 6207.r For
           more information about this cluster, see
           http://www.genoscope.cns.fr/
           cgi-bin/cluster.cgi?seq=CS0AAW2ZD03QP1&cluster=6207.r. Contact :
           Feng Liang Email : fliang@lifetech.com URL :
           http://fullength.invitrogen.com/ inVitroGen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0AAW2ZD03QP1.
FEATURES   Location/Qualifiers
     source   1..1200
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="CS0CAP003YH19"
           /tissue_type="THYMUS"
           /clone_lib="Homo sapiens THYMUS"
           /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
           with a NotI-oligo(dT) primer. Five prime end enriched,
           double-strand cDNA was digested with Not I and cloned into
           the Not I and EcoRV sites of the pCMVSPORT 6 vector.
           Library was not normalized."
ORIGIN

Query Match   2.0%;   Score 63.8;   DB 13;   Length 1200;
Best Local Similarity   47.6%;   Pred. No. 0.0083;   Indels 207;   Gaps   0;
Matches 188;   Conservative   0;   Mismatches 207;   Indels   0;

```
Qy   223 AACTTTATAACAGCAGCTAATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTA 282
              |||| |||  |||||   |||| |||||||
Db   806 AACTGTTTTATTTCACAAAATTTAAAATAAATTTATAAAATATAAAAAAAAAAAACA 865

Qy   283 TATGGTTTAGATCTAAGTAGTGATGATGATTGTGGTGGATACAAAGTAAACAAAGTAGTTCAGAT 342
              ||  ||| |||||||||
Db   866 AAAAATTTCAAATTTAATTTTTATCAAAAAATAAAAAATAAAAAAATTGTATAT 925

Qy   343 TACACAACTAGCAGAAATAGATTTGATCAAAGAAATAATTAATTTTTTATTGAATTAAA 402
              |||||||||||||||||
Db   926 GAAAAATAAAAATAATTTGAAAAATAAATAATTTTTTATTGAATTAAA 985

Qy   403 GTTAATGATGAAGCTAACGTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGT 462
              |||||||  ||||
Db   986 TTTAAAATAAAAAAATTTAAAAAATTAAAAAATAAAATTTGAAACAATAAAATTGGAA 1045

Qy   463 AATAGAACACAACAATCTAAGTTTGTAAGTTGGTGGTGTTGTGAAAATCCAGCTCACGTAATT 522
              ||||||  |||||||   ||||
Db  1046 ATTAAAAATTGATGATTAAAATTGGATAAATTGAAAAATTGAACATTGATAATTAAAAAATT 1105

Qy   523 AGATTACTGATGATGGGACTAAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTT 582
              ||   ||||  |||
Db  1106 GAAAATAATGATAATTGAAAAAAATGAAATAAAAAAAAATAAAATAAATTTTTT 1165

Qy   583 AATGACTTCATTTTAGATGATGCCGCCAATCTTACCTAA 617
              |||   |||||
Db  1166 AAATTTAAAATTTAATAACAATTATTTTTAAAAAA 1200
```

RESULT 10
FR0018463/c
LOCUS   FR0018463   470 bp   DNA   linear   GSS 09-DEC-1997
DEFINITION F.rubripes GSS sequence, clone 016E10aC6, genomic survey sequence.
ACCESSION   AL011359
VERSION   AL011359.1   GI:2676793
KEYWORDS   GSS; genome survey sequence.

FEATURES
    source
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
    source   Location/Qualifiers
            1..480
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMBMa011lJ17"
            /clone_lib="ZM_0.7-1.5_KB"
            /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
ORIGIN

Query Match   1.9%;   Score 61.8; DB 28;   Length 480;
Best Local Similarity 52.5%; Pred. No. 0.021;
Matches 135; Conservative   0; Mismatches 122; Indels   0; Gaps   0;

```
Qy  2923 GTTGATACCTTGACTGCTGCTGGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCT 2982
              ||  ||| |  | | ||||||   || |    ||| | | ||      || |  |||
Db    74 GCTACTACTACTACTGCTGCTACTACTACTGTACTACTACTAATGCTACTACTACTGCTGCT 133

Qy  2983 AACGTTAAGAAAAAACCTGCTGCTTAGGTGTCGGTGTAAATCGGTGGTGATAAGAAAACTGCT 3042
              ||  |  | | |  ||| ||| ||    |   |   ||     |||    ||||
Db   134 ACTACTACTACTGCTCCTACTACTGCTACTACTGCTTCTTCTACTACTACTACT 193

Qy  3043 GCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCT 3102
              |||  ||  ||| |||| |  |  |  ||  |      ||       |||  || ||
Db   194 GCTACTGCTACTGCTACTGCTACGACTACTACTCAGCTACTACTACTGCTACTCCT 253

Qy  3103 AAACCAACTGGGCTACTAAATCTGGTGCGCCTACAAAAAACCAACTGCTCCTCCTAAGCCAGCTGCT 3162
                || |   |  || | || |  | || |  |  ||| |  || |||   | ||  || |||
Db   254 GCTACTACTGCTACTACTACTGCTACTACTGCTACTACTACTGCTACTACTGCTACT 313

Qy  3163 CCAAAACCAACCGGCTCC 3179
                || ||| |  |||
Db   314 ACTGCTACTACTGCTAC 330
```

RESULT 12
BZ643413/c
LOCUS   BZ643413   843 bp   DNA   linear   GSS 29-JAN-2003
DEFINITION OGAOW57TM ZM_0.7_1.5_KB Zea mays genomic clone ZMBMa011lJ17,
            genomic survey sequence.
ACCESSION   BZ643413
VERSION   BZ643413.1 GI:28104915
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 843)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE   Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT   Other GSSs: OGAOW57TC
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES   Location/Qualifiers

    source
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMBMa011lJ17"
            /clone_lib="ZM_0.7-1.5_KB"
            /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
ORIGIN   1..843

Query Match   1.9%;   Score 61.8; DB 28;   Length 843;
Best Local Similarity 52.5%; Pred. No. 0.022;
Matches 135; Conservative   0; Mismatches 122; Indels   0; Gaps   0;

```
Qy  2923 GTTGATACCTTGACTGCTGTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCT 2982
              ||  ||| |  | ||||||   ||| |   ||| | | ||      || |  |||
Db   447 GCTACTACTACTGCTGCTGCTACTACTACTGTACTACTAATGCTACTACTACTGCTGCT 388

Qy  2983 AACGTTAAGAAAAAACCTGCTGCTTAGGTGTCGGTGTAAATCGGTGTAGATAGAAAACTGCT 3042
              ||  |  | | |  ||| ||| ||    |   |   ||    |||    ||||
Db   387 ACTACTACTACTGCTCCTACTACTGCTACTACTGCTTCTACTACTACTACT 328

Qy  3043 GCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTAGCTCACCAGCT 3102
              |||  ||  ||| |||| |  |  |  ||  |      ||       |||  || ||
Db   327 GCTACTGCTACTGCTACTACAGCTACTACGACTACTACTCAGCTACTACTACTGCTACTACTCCT 268

Qy  3103 AAACCAACTGGGCCTAAATCGGTGCGCCCTACAAAAACCAACTGCTCCTAGGCGCCAGCTGCT 3162
                || |   |  || | || |  | || |  |  ||| |  || |||   | ||  || |||
Db   267 GCTACTACTGCTACTACTACTGCTACTACTGCTACTACTACTGCTACTACTGCTACT 208

Qy  3163 CCAAAACCAACCGGCTCC 3179
                || ||| |  |||
Db   207 ACTGCTACTACTGCTAC 191
```

RESULT 13
CNS0161D/c
LOCUS   CNS0161D   1225 bp   DNA   linear   GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN15C18 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106171
VERSION   AL106171.1 GI:5620504
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1225)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeLoBAC11.
FEATURES   Location/Qualifiers
    source
            1..1225
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACN15C18"
            /clone_lib="DrosBAC"
            /plasmId="pBeloBAC11"
            /note="end : SP6"

ORIGIN

Query Match 1.9%; Score 61.8; DB 29; Length 1225;
Best Local Similarity 31.5%; Pred. No. 0.022;
Matches 129; Conservative 86; Mismatches 194; Indels 0; Gaps 0;

```
Qy   204 TCTTAGAGATGTTGATGATAACTTTATACACAGCAGCTAATGGAACAATTATCAAATTAGA 263
             |:|:| .|: ..:|| ::||:|| |:|:|:|||
Db  1213 TTTTTWTKATTTTTWWAWAWAWWWAWAAAAWWTTWTTATAAAAATATATAWATWTA 1154

Qy   264 TAGTTTTACTAAACCATTATATGGTTTAGATCTAAGTGATGATTGTGGTGGATACAAAGT 323
          |:||| .:|: ||||| |::|. |:|:|
Db  1153 TTATWWAAAWATTTTTTATATWATTWTWATAAAWAAAWAAAATATTAWWAWWWAAWAAW 1094

Qy   324 AAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGATTTGATCAAAGACAAACAAG 383
          :|:|. |:| .| |||| |:| ||||||
Db  1093 AWAWAAWNWATWWTWTWFKAATATAAAARTAWATATTAAAAATWAWAWWWTWAWAAAATAT 1034

Qy   384 AGCATATTATGCTCGTTGGTTAATGATGAAGCTAACGTTCATTTAAAAAGAATTAATAC 443
           |:| |: |: .:||
Db  1033 ATTTTTTTTTWTTTTAWAAAAAAWAAAWWATARATTATWWATWATRAAAWWWTARAAAAAAT 974

Qy   444 TAACTCAAATAGATTGGTAATAGAAACAACAATTCTAAGTTTGTAATTGGTGGTGTTGA 503
           |:||:| .::
Db   973 AAWTTATTTWAAAWTTTTTTRAAAWWTTTTTTTWWTTTAAAAWTWWTTATWTRTT 914

Qy   504 TAATCCAGCTCACGTAATTAGATTTACTGATCGATGATGGGGACTAAATTTAATTTACAAACCA 563
          :| :| ||||
Db   913 TWTRTRRARAAATWWWTTTTTTTTTTTTTTATATTTTTTTMYHHTYTYWTTHHGHM 854

Qy   564 AACTCAAGGTCGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCTTA 612
          ::: .::|  ||||
Db   853 MMMMWMHHMMWGBGBGKGGGTTTTTKKTTTKTTKTTTGWGATTTTA 805
```

RESULT 14
CNS04NSM/c
LOCUS   CNS04NSM 735 bp DNA linear GSS 01-SEP-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
          123M05 of library G from Tetraodon nigroviridis, genomic survey
          sequence.
ACCESSION AL299119
VERSION   AL299119.1 GI:8038260
KEYWORDS   GSS; genome survey sequence.
SOURCE   Tetraodon nigroviridis
ORGANISM   Tetraodon nigroviridis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
           Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
           Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
           Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
           Saurin,W. and Weissenbach,J.
TITLE   Estimate of human gene number provided by genome-wide analysis
           using Tetraodon nigroviridis DNA sequence
JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE   20296633
PUBMED   10835645
REFERENCE 2
AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
           Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
           Saurin,W., Bernot,A. and Weissenbach,J.
TITLE   Characterization and repeat analysis of the compact genome of the
           freshwater pufferfish Tetraodon nigroviridis
JOURNAL   Genome Res. 10 (7), 939-949 (2000)
MEDLINE   20359837
PUBMED   10891143
REFERENCE 3 (bases 1 to 735)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT   - Web : www.genoscope.cns.fr)
           This sequence is a single read and was generated as part of a large
           scale clone-end sequencing project of the Tetraodon nigroviridis
           genome. For more information, Please take a look at
           http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
     source   1..735
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="123M05"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : COBG123AG03LP1-end : T7"

ORIGIN

Query Match 1.9%; Score 61.6; DB 29; Length 735;
Best Local Similarity 50.0%; Pred. No. 0.024;
Matches 130; Conservative 8; Mismatches 122; Indels 0; Gaps 0;

```
Qy  2922 AGTTGATACCTTGACTGCGTCGTGGTTCAGTTTACAAGAGAGATTATTACCCAAACTGC 2981
              :||::  :|:.
Db   500 AGCTRCTACTACTGCTACWATTKCWGCTGCTACTACTACTACTACTAC 441

Qy  2982 TAACGTTAAGAAAAAAAACCTGCTTTAGGTGCTGGTAAATCTGGTGATGATAAGAAACCGC 3041
            ||||  :| |||| |||| ||| |:|
Db   440 TACTGCTACTACTGCTGCTACTACTGCTCTACTACTACTGCTGC 381

Qy  3042 TGCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGC 3101
           ||| ||| | ||| |||||
Db   380 TGCTACTACKGCKGCTGCTACTACTACTACTAC 321

Qy  3102 TAAACCAACTGGGGCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGC 3161
           |||||  :| ||||
Db   320 TACTGCTACTACTRCTACTACTACTACTACTACTACTAC 261

Qy  3162 TCCAAAACCAACCGCTCCCA 3181
           |||||
Db   260 TACTACTACTACTACTA 241
```

RESULT 15
AQ946120/c
LOCUS   AQ946120 641 bp DNA linear GSS 27-JAN-2000
DEFINITION Sheared DNA-46J23.TR Sheared DNA Trypanosoma brucei genomic clone
           Sheared DNA-46J23, genomic survey sequence.
ACCESSION AQ946120
VERSION   AQ946120.1 GI:6769385
KEYWORDS   GSS.
SOURCE   Trypanosoma brucei
ORGANISM   Trypanosoma brucei
           Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
           Trypanosoma.
REFERENCE 1 (bases 1 to 641)
AUTHORS   El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
           Gerrard,C., Leech,V., de Jong,P., Ulu,E., Melville,S.,
           Donelson,J., Fraser,C. and Adams,M.
TITLE   Determination of clone end sequences from Trypanosoma brucei GUTat
           10.1 sheared DNA library
JOURNAL   Unpublished (1999)
COMMENT   Other GSSs: Sheared DNA-46J23.TF
           Contact: Najib M. El-Sayed
           Department of Eukaryotic Genomics
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: nelsayed@tigr.org
           Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
           DNA library constructed at TIGR. Clones will be available for
           distribution through ATCC. Sheared DNA end sequences search page:
           http://www.tigr.org/tdb/mdb/tbdb/.
           Seq primer: M13-Reverse
           Class: shotgun.

FEATURES   Location/Qualifiers
source   1..641
                 /organism="Trypanosoma brucei"
                 /mol_type="genomic DNA"
                 /strain="TREU927/4 GUTat 10.1"
                 /db_xref="taxon:5691"
                 /clone="Sheared DNA-46J23"
                 /clone_lib="Sheared DNA"
                 /note=Vector: pUC18; Site_1: SmaI; Constructed at The
                 Institute for Genomic Research (TIGR), Rockville, MD.
                 Genomic DNA isolated from a cloned population of
                 Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                 sheared to give a tight size distribution (approx 2 kb).
                 The v + i method used for the library construction is
                 described in detail in Smith, H.O. and Venter, J.C.
                 (Making small insert libraries for whole genome shotgun
                 sequencing projects. In Genome Sequencing: A Practical
                 Approach, eds. M. Vaudin and B. Barell, Oxford University
                 Press, 1999)."

ORIGIN

Query Match   1.9%;   Score 60.6; DB 28; Length 641;
Best Local Similarity   44.5%;   Pred. No. 0.038;
Matches 240; Conservative   0; Mismatches 299; Indels   0; Gaps   0;

Qy   436 ATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTTGTAATTGGT 495
          || ||| |||||| | || ||| |||| ||||||||||||||||||||||||
Db   617 AGTACTACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 558

Qy   496 GGTGTTGATAATCCAGCTCACGTAATTAGAATTTACTGATGATGGGACTAAATTTAATTTT 555
          |||| | || |||||||||||||||||||||||||||||||||||||| |||| |||
Db   557 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 498

Qy   556 ACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCTTACCT 615
          ||| |||||||||||||||||||| || ||||||||||||||||||||||||||||
Db   497 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 438

Qy   616 AAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAGATCTTACCAAATGAC 675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 378

Qy   676 GTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGAACAAAATGCTGATGAT 735
          | ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   377 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 318

Qy   736 GGGATGTTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCTATTGCTCAAACTAAA 795
          ||||| |||| | || || ||||| ||||| || |||||||||
Db   317 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 258

Qy   796 ACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGCCTGGTGCAAAC 855
          || || | | | || ||||||||||||||||
Db   257 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 198

Qy   856 AATAGATACGATCTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTTTCCAATTAGAT 915
          || || ||| ||||||| || | ||| |||||||||||||||
Db   197 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 138

Qy   916 GAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAAAATATTACAAGATT 974
          |||| ||||||||||||||||||||||||||||||||||
Db   137 AATAATAATAATAATAATAATAATAATGATGATGATGATAATAATAATAGAAGAGT 79

Search completed: May 5, 2004, 19:58:28
Job time : 7298.65 secs